

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP2005/015770

International filing date: 30 August 2005 (30.08.2005)

Document type: Certified copy of priority document

Document details: Country/Office: JP
Number: 2004-259725
Filing date: 07 September 2004 (07.09.2004)

Date of receipt at the International Bureau: 06 October 2005 (06.10.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 4 年 9 月 7 日

出 願 番 号
Application Number: 特 願 2 0 0 4 - 2 5 9 7 2 5

パリ条約による外国への出願
に用いる優先権の主張の基礎
となる出願の国コードと出願
番号
J P 2 0 0 4 - 2 5 9 7 2 5
The country code and number
of your priority application,
to be used for filing abroad
under the Paris Convention, is

出 願 人
Applicant(s): インターマン株式会社

2 0 0 5 年 9 月 2 1 日

特許庁長官
Commissioner,
Japan Patent Office

中 嶋



【書類名】	特許願
【整理番号】	Z096-0001
【提出日】	平成16年 9月 7日
【あて先】	特許庁長官殿
【国際特許分類】	G06F 17/60
【発明者】	
【住所又は居所】	鹿児島県鹿児島市上荒田町 1 3 番 2 2 号
【氏名】	上田平 重樹
【発明者】	
【住所又は居所】	鹿児島県鹿児島市上荒田町 1 3 番 2 2 号
【氏名】	上田平 美嗣
【特許出願人】	
【識別番号】	596177559
【氏名又は名称】	インターマン株式会社
【代理人】	
【識別番号】	100105924
【弁理士】	
【氏名又は名称】	森下 賢樹
【電話番号】	03-3461-3687
【手数料の表示】	
【予納台帳番号】	091329
【納付金額】	16,000円
【提出物件の目録】	
【物件名】	特許請求の範囲 1
【物件名】	明細書 1
【物件名】	図面 1
【物件名】	要約書 1

【書類名】 特許請求の範囲

【請求項 1】

検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、

ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、

前記収集部による前記評価対象のコンテンツの収集の可否に応じて、前記評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、

を備えることを特徴とする情報検索提供装置。

【請求項 2】

ユーザから検索条件文を取得する検索条件文取得部と、

前記取得した検索条件文に合致したコンテンツを検索する検索部と、

前記検索されたコンテンツの前記信頼度をユーザに提示する信頼度提示部と、

をさらに備えることを特徴とする請求項 1 に記載の情報検索提供装置。

【請求項 3】

前記評価対象のコンテンツに含まれる当該コンテンツの存続期間の推測に供する要素に基づいて、前記評価対象のコンテンツの信頼度を予測する予測部をさらに備え、

前記信頼度提示部は前記予測された信頼度をユーザに提示することを特徴とする請求項 2 に記載の情報検索提供装置。

【請求項 4】

前記信頼度制御部は、前記収集部により前記評価対象のコンテンツが所定のタイミングで連続して収集される期間内において、前記評価対象のコンテンツの信頼度を徐々に増加せしめることを特徴とする請求項 1 から 3 のいずれかに記載の情報検索提供装置。

【請求項 5】

前記収集部による前記評価対象の収集の可否の履歴に応じて当該評価対象のコンテンツの信頼度の増加幅を制御する履歴増加幅制御部をさらに備えることを特徴とする請求項 1 から 4 のいずれかに記載の情報検索提供装置。

【請求項 6】

前記評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断する関連性判断部をさらに備え、

前記信頼度制御部は、前記関連性判断部により関連性があると判断されたときにおける前記評価対象のコンテンツの信頼度の増加幅を、関連性がないと判断されたときにおける前記評価対象のコンテンツの信頼度の増加幅よりも高く制御することを特徴とする請求項 1 から 5 のいずれかに記載の情報検索提供装置。

【請求項 7】

前記関連性判断部は、前記信頼度の高いコンテンツ内に、当該コンテンツと前記評価対象のコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて、前記評価対象のコンテンツと前記信頼度の高いコンテンツとの関連性の有無を判断することを特徴とする請求項 6 に記載の情報検索提供装置。

【請求項 8】

検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、

ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、

前記収集された評価対象のコンテンツの内容と、当該コンテンツを掲載するウェブページと同一のウェブページから次に収集される評価対象のコンテンツの内容とを比較する比較部と、

前記比較の結果に基づいて前記評価対象のコンテンツが更新されたか否かを判断する更新判断部と、

前記判断の結果に応じて前記評価対象のコンテンツの信頼度を増減せしめる信頼度制御

部と、
を備えることを特徴とする情報検索提供装置。

【請求項 9】

前記信頼度制御部は、前記更新判断部により前記評価対象のコンテンツが更新されていると所定のタイミングで連続して判断される期間内において、前記評価対象のコンテンツの信頼度を徐々に増加せしめることを特徴とする請求項 8 に記載の情報検索提供装置。

【請求項 10】

前記収集された評価対象のコンテンツの内容を参照し当該コンテンツが本来更新を必要とするべきものであるか否かを分析する更新分析部をさらに備え、

前記信頼度制御部は、前記更新判断部による判断の結果および前記更新分析部による分析の結果に応じて、前記評価対象のコンテンツの信頼度を増減せしめることを特徴とする請求項 8 に記載の情報検索提供装置。

【請求項 11】

検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、

評価対象のコンテンツに関連するコンテンツを取得する関連コンテンツ取得部と、

前記取得されたコンテンツを参照し前記評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出する要素抽出部と、

前記抽出された要素を参照して前記評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、

を備えることを特徴とする情報検索提供装置。

【請求項 12】

前記要素抽出部は、前記信頼度の評価の観点ごとに、前記評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出するものであり、前記信頼度制御部は、前記信頼度の評価の観点ごとに抽出された要素を参照して、前記評価対象のコンテンツの信頼度を増減せしめることを特徴とする請求項 11 に記載の情報検索提供装置。

【請求項 13】

前記抽出された要素それぞれを、前記評価対象のコンテンツの信頼度の増加に貢献する第 1 グループ、あるいは前記評価対象のコンテンツの信頼度の減少に貢献する第 2 グループに分類する分類部をさらに備え、

前記信頼度制御部は、前記分類部により分類された前記第 1 グループ内の要素数が前記第 2 グループ内の要素数よりも多いとき、前記評価対象のコンテンツの信頼度を増加せしめることを特徴とする請求項 11 または 12 に記載の情報検索提供装置。

【請求項 14】

検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、

判断対象項目を含むコンテンツを取得するコンテンツ取得部と、

前記取得されたコンテンツのうち、前記判断対象項目の情報発生源となる自コンテンツを特定する自コンテンツ特定部と、

前記特定された自コンテンツの内容と、前記取得されたコンテンツのうち前記自コンテンツ以外の他コンテンツの内容とを比較し、前記判断対象項目の整合性の有無を判断する整合性判断部と、

前記判断の結果に応じて前記自コンテンツの信頼度と前記他コンテンツの信頼度とを増減せしめる信頼度制御部と、

を備えることを特徴とする情報検索提供装置。

【請求項 15】

前記自コンテンツの信頼度と前記他コンテンツの信頼度とがともに高いときであって、前記整合性判断部により前記判断対象項目について整合性有りと判断されたとき、前記信頼度制御部は前記自コンテンツの信頼度と前記他コンテンツの信頼度とをともに増加せしめることを特徴とする請求項 14 に記載の情報検索提供装置。

【請求項 1 6】

ウェブページを公開するサーバの所在を示す所在情報を参照して当該サーバにアクセスするアクセス部と、

前記アクセス部による前記サーバへのアクセスの可否を判断するアクセス可否判断部と、

前記アクセス可否判断部による判断の結果と前記アクセス部により前記サーバにアクセスしたタイミングとの対応関係を保持するアクセス履歴保持部と、

前記対応関係を参照し前記アクセス部による前記サーバへのアクセスが最初に可能になったタイミングを特定するタイミング特定部と、

をさらに備え、

前記信頼度制御部は前記特定されたタイミングにて前記信頼度の増減の制御を開始することを特徴とする請求項 1 から 1 5 のいずれかに記載の情報検索提供装置。

【請求項 1 7】

ネットワーク上に接続される端末装置と、前記ネットワークを介してウェブページを提供するサーバと、当該サーバ上のウェブページに掲載されるコンテンツを収集し、前記収集されたコンテンツのうち検索条件文に合致したコンテンツを検索結果として前記端末装置に提供する情報検索提供装置と、を含む情報検索提供システムであって、

前記情報検索提供装置は、

ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、

前記収集部による前記評価対象のコンテンツの収集の可否に応じて、前記評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、
を備えることを特徴とする情報検索提供システム。

【書類名】 明細書

【発明の名称】 情報検索提供装置および情報検索提供システム

【技術分野】

【０００１】

この発明は情報検索提供装置および情報検索提供システムに係るものであり、特にインターネットなどのネットワーク上のウェブページに掲載される、文章、画像やプログラムなどのコンテンツを検索結果として提供する情報検索提供装置および情報検索提供システムに関する。

【背景技術】

【０００２】

近年、インターネットの利用が進むにつれ、その主要なアプリケーションであるWWW（World Wide Web）のユーザが激増し、電子化された情報の流通が非常に活発になってきた。それに伴い、情報量の爆発的な増大によってユーザが真に欲する情報を取り出すことが困難になるという問題が生じてきた。このような問題を軽減するために、インターネットなどのネットワーク上において様々な検索エンジンが登場してきた。

【０００３】

そうした様々な検索エンジンの使い勝手をよくするために、例えば、特許文献１には、過去にユーザが検索を実行し、どのウェブページで目的の情報を見出されたかを推定し、その推定結果から検索キーワードおよびウェブページのURL（Uniform Resource Locator）にポイントを付与し、このポイントの大きい順に、後の検索におけるコンテンツの検索結果としての提示順序を決定する技術が開示されている。

【特許文献１】 特開２００４－２９９４３号公報

【発明の開示】

【発明が解決しようとする課題】

【０００４】

確かに、特許文献１によれば、所望の情報が含まれるウェブページがユーザにより見出される可能性は高くなるが、ユーザがウェブページに掲載されているコンテンツを利用したとき、そのコンテンツが信頼のおけないコンテンツであれば、不利益を被る可能性が高い。そのため、そのような不都合を軽減する新たな情報検索サービスの仕組み作りが必要である。

【０００５】

本発明はこうした課題に鑑みてなされたものであり、その目的は信頼におけるコンテンツと信頼のおけないコンテンツとを選別する新たな情報検索サービスを実現する情報検索提供装置および情報検索提供システムの提供にある。

【課題を解決するための手段】

【０００６】

本発明のある態様は、情報検索提供装置に関する。この情報検索提供装置は、検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、収集部による評価対象のコンテンツの収集の可否に応じて、評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、を備える。

【０００７】

この態様によれば、評価対象のコンテンツの収集の可否に応じて当該コンテンツを信頼度の側面から評価することができ、その結果、信頼におけるコンテンツと信頼のおけないコンテンツとを選別でき、ユーザ利便性の高い情報検索サービスを実現できる。

【０００８】

この態様の情報検索提供装置は、ユーザから検索条件文を受け付ける検索条件文受付部と、受け付けた検索条件文に合致したコンテンツを検索する検索部と、検索されたコンテンツの信頼度をユーザに提示する信頼度提示部と、をさらに備えてもよい。評価対象のコンテンツに含まれる当該コンテンツの存続期間の推測に供する要素に基づいて、評価対象

のコンテンツの信頼度を予測する予測部をさらに備え、信頼度提示部は予測された信頼度をユーザに提示してもよい。

【0009】

信頼度制御部は、収集部により評価対象のコンテンツが所定のタイミングで連続して収集される期間内において、前記評価対象のコンテンツの信頼度を徐々に増加せしめてもよい。例えば、第1のタイミング以降、前記収集部により前記評価対象のコンテンツが収集され、その後、第2のタイミング以降、前記収集部により前記評価対象のコンテンツが収集されず、さらにその後、第3のタイミング以降、前記収集部により前記評価対象のコンテンツが収集されたとき、前記増加幅制御部は、前記第3のタイミングにおける前記評価対象のコンテンツの信頼度の増加幅を、前記第1のタイミングにおける前記評価対象のコンテンツの信頼度の増加幅よりも低く制御してもよい。

【0010】

この態様の情報検索提供装置は、収集部による評価対象の収集の可否の履歴に応じて当該評価対象のコンテンツの信頼度の増加幅を制御する履歴増加幅制御部をさらに備えてもよい。評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断する関連性判断部をさらに備え、信頼度制御部は、関連性判断部により関連性があると判断されたときにおける評価対象のコンテンツの信頼度の増加幅を、関連性がないと判断されたときにおける評価対象のコンテンツの信頼度の増加幅よりも高く制御してもよい。

【0011】

関連性判断部は、評価対象のコンテンツと信頼度の高いコンテンツとの間に関連性があることを示す関連情報が、評価対象のコンテンツ内に含まれるか否かに基づいて、評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断してもよい。関連性判断部は、信頼度の高いコンテンツ内に、当該コンテンツと評価対象のコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて、評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断してもよい。

【0012】

本発明の別の態様も、情報検索提供装置に関する。この情報検索提供装置は、検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、収集された評価対象のコンテンツの内容と、当該コンテンツを掲載するウェブページと同一のウェブページから次に収集される評価対象のコンテンツの内容とを比較する比較部と、比較の結果に基づいて評価対象のコンテンツが更新されたか否かを判断する更新判断部と、判断の結果に応じて評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、を備える。

【0013】

信頼度制御部は、更新判断部により評価対象のコンテンツが更新されていると所定のタイミングで連続して判断される期間内において、評価対象のコンテンツの信頼度を徐々に増加せしめてもよい。収集された評価対象のコンテンツの内容を参照し当該コンテンツが本来更新を必要とするべきものであるか否かを分析する更新分析部をさらに備え、信頼度制御部は、更新判断部による判断の結果および前記更新分析部による分析の結果に応じて、評価対象のコンテンツの信頼度を増減せしめてもよい。

【0014】

本発明のさらに別の態様も、情報検索提供装置に関する。この情報検索提供装置は、検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、評価対象のコンテンツに関連するコンテンツを取得する関連コンテンツ取得部と、取得されたコンテンツを参照し評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出する要素抽出部と、抽出された要素を参照して評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、を備える。

【0015】

要素抽出部は、信頼度の評価の観点ごとに、評価対象のコンテンツの信頼度の増加ある

いは減少に貢献する要素を抽出するものであり、信頼度制御部は、信頼度の評価の観点ごとに抽出された要素を参照して、評価対象のコンテンツの信頼度を増減せしめてもよい。

【0016】

この態様の情報検索提供装置は、抽出された要素それぞれを、評価対象のコンテンツの信頼度の増加に貢献する第1グループ、あるいは評価対象のコンテンツの信頼度の減少に貢献する第2グループに分類する分類部をさらに備え、信頼度制御部は、分類部により分類された第1グループ内の要素数が第2グループ内の要素数よりも多いとき、評価対象のコンテンツの信頼度を増加せしめてもよい。

【0017】

本発明のさらに別の態様も、情報検索提供装置に関する。この情報検索提供装置は、検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置であって、判断対象項目を含むコンテンツを取得するコンテンツ取得部と、取得されたコンテンツのうち判断対象項目の情報発生源となる自コンテンツを特定するコンテンツ特定部と、特定された自コンテンツの内容と取得されたコンテンツのうち自コンテンツ以外の他コンテンツの内容とを比較し、判断対象項目の整合性の有無を判断する整合性判断部と、判断の結果に応じて自コンテンツの信頼度と他コンテンツの信頼度とを増減せしめる信頼度制御部と、を備える。

【0018】

この態様の情報検索提供装置は、自コンテンツの信頼度と他コンテンツの信頼度とがともに高いときであって、整合性判断部により判断対象項目について整合性有りと判断されたとき、信頼度制御部は自コンテンツの信頼度と他コンテンツの信頼度とをともに増加せしめてもよい。

【0019】

上述の情報検索提供装置は、ウェブページを公開するサーバの所在を示す所在情報を参照して当該サーバにアクセスするアクセス部と、アクセス部によるサーバへのアクセスの可否を判断するアクセス可否判断部と、アクセス可否判断部による判断の結果とアクセス部によりサーバにアクセスしたタイミングとの対応関係を保持するアクセス履歴保持部と、対応関係を参照しアクセス部によるサーバへのアクセスが最初に可能になったタイミングを特定するタイミング特定部と、をさらに備え、信頼度制御部は特定されたタイミングにて信頼度の増減の制御を開始してもよい。

【0020】

本発明のさらに別の態様は、情報検索提供システムに関する。この情報検索提供システムは、ネットワーク上に接続される端末装置と、ネットワークを介してウェブページを提供するサーバと、当該サーバ上のウェブページに掲載されるコンテンツを収集し、収集されたコンテンツのうち検索条件文に合致したコンテンツを検索結果として端末装置に提供する情報検索提供装置と、を含む情報検索提供システムであって、情報検索提供装置は、ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部と、収集部による評価対象のコンテンツの収集の可否に応じて、評価対象のコンテンツの信頼度を増減せしめる信頼度制御部と、を備える。

【0021】

本発明のさらに別の態様は、情報検索提供方法に関する。この情報検索提供方法は、ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集するステップと、評価対象のコンテンツの収集の可否に応じて、評価対象のコンテンツの信頼度を増減せしめるステップと、ユーザから検索条件文を取得するステップと、収集されたコンテンツのうち取得された検索条件文に合致したコンテンツを検索結果として提供するステップと、を有する。

【発明の効果】

【0022】

本発明によれば、ユーザ利便性の高い情報検索サービスを提供できる。

【発明を実施するための最良の形態】

【0023】

実施の形態1

図1は、本実施の形態に係る情報検索提供システム10の構成を示す。図1に示す情報検索提供システム10は、情報検索提供装置100と、ウェブサーバ200と、ネットワーク12とを備える。ネットワーク12には、情報検索提供装置100と、ウェブサーバ200と、端末装置50とが接続されている。本実施の形態に係るネットワーク12はWAN（Wide Area Network）であるが、別の例として、例えばLAN（Local Area Network）、あるいは所定のデータをTCP/IP（Transmission Control Protocol/Internet Protocol）等の通信規格に準じて送受信する通信網であってよい。端末装置50は、パーソナルコンピュータ、PDA（Personal Digital Assistance）、あるいは携帯電話であり、一方、情報検索提供装置100やウェブサーバ200はパーソナルコンピュータなどで実装される。ウェブサーバ200はウェブページに掲載したコンテンツをネットワーク12を介して公開する。コンテンツとはテキストデータ、映像データ、動画データ、音声データなど電子化可能なデータを指す。

【0024】

ユーザは、端末装置50に搭載される図示しないWWWブラウザを用いて情報検索提供装置100にアクセスし、検索条件文を指定して検索を行う。情報検索提供装置100は、その検索条件文による検索実行指示を受けて、図1では図示しないデータベースに保持されている膨大な量の情報の中から、検索条件文に合致したコンテンツと当該コンテンツに付随する情報（以下、「コンテンツ付随情報」という）とを抽出し、それらを一覧化したものを検索結果として端末装置50に提供する。このとき、さらにコンテンツの信頼度が提示される。コンテンツ付随情報はウェブページのURL、コンテンツのサイズ、コンテンツの更新日付、コンテンツの表題などを含む。なお、データベースに保持されている膨大な量のコンテンツは、情報検索提供装置100内の図1では図示しない収集制御部によりウェブページから定期的あるいは不定期に収集される。

【0025】

図2は、本実施の形態に係る端末装置50の構成を示す。端末装置50は、検索条件文受付部62と、送信部52と、取得部54と、表示部64とを備える。検索条件文受付部62は、キーボード、マウス、その他の外部機器を介してユーザから入力されたデータ、ここでは検索条件文を取得する。検索条件文は、例えば「牛肉」、「料理」、「レシピ」のようなキーワード単位の形式でもよいし、「牛肉を使った料理のレシピが知りたい」というように自然文の形式でもよい。

【0026】

送信部52は検索条件文受付部62により取得されたユーザからの検索条件文をネットワーク12を介して情報検索提供装置100に送信する。このとき、情報検索提供装置100による検索処理が行われる。取得部54は、検索結果として、検索条件文に合致したコンテンツやコンテンツ付随情報、例えば当該コンテンツのサイズやURLの一覧をネットワーク12を介して取得する。このとき、さらに、コンテンツの信頼度が提示される。表示部64は取得部54により取得された情報を整形し、図示しないディスプレイを介してユーザに表示する。なお、図示しないスピーカを介して音声によりユーザに通知してもよい。

【0027】

図3は、本実施の形態に係る情報検索提供装置100の構成を示す。情報検索提供装置100は、ユーザから指定された検索条件文に合致したコンテンツを検索し検索結果として端末装置50に提供する検索制御部110と、ウェブサーバ200上のウェブページに掲載されるコンテンツを定期的あるいは不定期に収集する収集制御部120と、収集されたコンテンツを保持する検索情報保持部180と、収集制御部120によるウェブサーバへのアクセス履歴を保持するアクセス履歴保持部190とを備える。検索制御部110は、検索条件文取得部112と、検索部114と、検索結果提供部116と、信頼度提示部118とを備える。収集制御部120は、アクセス部122と、アクセス可否判断部12

４と、タイミング特定部１２６と、収集部１２８と、信頼度制御部１３０と、履歴増加幅制御部１３２とを備える。

【００２８】

情報検索提供装置１００の各構成要素は、ハードウェアコンポーネントでいえば、任意のコンピュータのＣＰＵ（Central Processing Unit）、メモリ、その他の素子、およびソフトウェアとしてメモリにロードされた情報検索、収集機能、記憶機能、およびその他機能のあるプログラムなどによって実現されるが、ここではそれらの連携によって実現される機能ブロックを描いている。したがって、これらの機能ブロックがハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現できることは、当業者には理解されるところである。

【００２９】

検索条件文取得部１１２は送信部５２により送信されたユーザからの検索条件文を取得する。このとき、検索条件文が自然文の形式であれば、検索条件文取得部１１２はその検索条件文を一度、図示しないキーワード抽出部に送ってもよい。このとき、キーワード抽出部は、形態素へ分解し、主に名詞、先の検索条件文の例で言えば、「牛肉」、「料理」、「レシピ」などをキーワードとして抽出し、検索条件文取得部１１２は、それらキーワードを組み合わせたものを新たな検索条件文として取得する。

【００３０】

検索部１１４は、検索情報保持部１８０に格納されている検索情報ファイルから、検索条件文取得部１１２で取得された検索条件文に合致したコンテンツ、当該コンテンツの信頼度やコンテンツ付随情報を検索する。このとき、検索部１１４は、検索情報保持部１８０に格納されているコンテンツに対する索引語を含むファイル、いわゆるインデックスファイルを利用して検索を行ってもよい。これにより検索処理の高速化を実現できる。検索結果提供部１１６は、検索により得られた情報を一覧化したものを検索結果として端末装置５０に提供する。このとき、信頼度提示部１１８は検索されたコンテンツの後述する信頼度を端末装置５０に提示する。これにより、ユーザはコンテンツがどの程度の信頼度を有するかを知ることができる。

【００３１】

アクセス部１２２は、後述するアクセス履歴保持部１９０内のアクセス履歴ファイルに含まれるウェブサーバ２００の所在を示す所在情報、例えばＩＰアドレスを参照して、ウェブサーバ２００にアクセスする。別の例として、アクセス部１２２はウェブページのＵＲＬやドメイン名を参照して、当該ウェブページにアクセスしてもよい。本実施の形態に係るアクセス部１２２は、アクセス履歴ファイルに含まれる、未使用のＩＰアドレスを含むすべてのＩＰアドレスに対し、一日に一回、例えば午前０時に順番にアクセスするものとする。なお、アクセス部１２２によるアクセスの頻度は「一日に一回」に限定されるものではなく、「一日に二回」でも「二日に一回」でもよい。アクセス部１２２は定期的ではなく不定期にアクセスしてもよい。

【００３２】

アクセス可否判断部１２４は、アクセス部１２２によるウェブページへのアクセスの可否を判断する。このとき、アクセス可否判断部１２４は、アクセスが不可能であったと判断したとき、当該ＩＰアドレスでかつ後述のアクセス履歴欄内の現在の日付の欄にフラグ「×」を設定する。一方、アクセス可否判断部１２４は、アクセス部１２２によるアクセスが可能であったと判断したとき、当該ＩＰアドレスでかつ後述のアクセス履歴欄内の現在の日付の欄にフラグ「○」を設定する。

【００３３】

タイミング特定部１２６はアクセス履歴ファイルを参照しアクセス部１２２によるウェブページへのアクセスが最初に可能になったタイミングを特定する。具体的には、タイミング特定部１２６は、過去のアクセス履歴が今まで「×」であったが、最初に「○」になった日をウェブページの提供開始日とする。タイミング特定部１２６は現在の日付がウェブページの提供開始日に相当すると判断したとき、後述のウェブページ提供開始日欄に現

在の日付を設定する。すでにウェブページがネットワーク 12 上に存在し、以後にアクセス部 122 によるアクセスがあったとき、タイミング特定部 126 はアクセスがあったときの日付をウェブページの提供開始日に相当すると判断してもよい。

【0034】

ところで、近年、インターネットに接続されるサーバの増加に伴い、IP アドレスの枯渇化が問題になっている。本実施の形態によれば、長期間「×」である IP アドレスを未使用とすることで、未使用の IP アドレスを特定できる。その結果、IP アドレスを管理する第三者機関などが、その未使用の IP アドレスを他の IP アドレスの取得を望む情報提供者に提供することで、効果的に IP アドレスを活用できる。

【0035】

アクセス履歴保持部 190 は、アクセス可否判断部 124 による判断の結果とアクセス部 122 によりウェブページにアクセスしたタイミングとの対応関係を含むアクセス履歴ファイルを保持する。アクセス履歴ファイルには、ネットワーク 12 上に存在する可能な限りすべての IP アドレスが保持されているものとする。もちろん、すべてではなく一部の IP アドレスが保持されていてもよい。検索情報保持部 180 は後述する収集部 128 により収集されたコンテンツ、コンテンツ付随情報、信頼度などを含む検索情報ファイルを保持する。

【0036】

収集部 128 は、アクセス可否判断部 124 によりウェブページへのアクセスが可能であったと判断されたとき、そのウェブページに掲載される評価対象のコンテンツを所定のタイミングで、ここでは一日に一回収集する。なお、アクセス部 122 により IP アドレスやドメイン名を利用したアクセスがあったときは、収集部 128 は、ウェブサーバ上のすべてのウェブページが掲載するコンテンツを収集してもよい。コンテンツ収集の際、収集部 128 は収集されたコンテンツのコンテンツ付随情報を生成し、検索情報保持部 180 に格納する。収集部 128 は、コンテンツ内の HTML (Hyper Text Markup Language) 文の先頭箇所、例えば<HEAD>～</HEAD>間にメタタグが記載されていれば、このメタタグで囲まれる領域内に記述された指示内容に従ってもよい。すなわち、情報検索提供装置 100 への登録の不要が明示されていれば収集部 128 は当該ページの情報を収集しなくてもよい。

【0037】

信頼度制御部 130 は収集部 128 による評価対象のコンテンツの収集の可否に応じて、評価対象のコンテンツの信頼度を所定のタイミングで、ここでは一日に一回増減せしめる。具体的には、信頼度制御部 130 は、評価対象のコンテンツをウェブページから収集できたとき、当該評価対象のコンテンツに設定された信頼度を所定の増加幅分だけ増加させる。一方、収集部 128 により評価対象のコンテンツをウェブページから収集できなかったとき、当該評価対象のコンテンツに設定された信頼度を減少させる。なお、信頼度制御部 130 はウェブページの提供開始日から、そのウェブページに掲載されるコンテンツの信頼度の増減の制御を開始する。

【0038】

さらに、信頼度制御部 130 は、収集部 128 により評価対象のコンテンツが所定のタイミングで連続して収集される期間内において、評価対象のコンテンツの信頼度を徐々に増加させる。なお、本実施の形態では、コンテンツの信頼度を増減させているが、これは実質的にコンテンツを掲載するウェブページの信頼度、当該ウェブページを公開するウェブサーバや当該ウェブサーバを運営する情報提供者の信頼度を増減させることと同意である。さらに、当該ウェブサーバを運営する情報提供者が提供するコンピュータシステムに含まれる、他の各種サービス提供サーバ、例えば、メールサーバやFTP (File Transfer Protocol) サーバやニュースサーバなどの信頼度を増減させることと実質的に同意である。

【0039】

本実施の形態に係る情報検索提供装置 100 によれば、コンテンツに信頼度を設定する

ことで、そのコンテンツがどの程度の信頼度を有するのかをユーザに提示できる。その結果、ユーザは、情報検索提供装置 100 が提供する情報検索サービスを通じて、コンテンツの信頼度を手軽に知ることができる。例えば、ある会社が他の企業と取引するとき、相手の素性を知りたいときがある。このとき、その企業が提供するウェブページの信頼度を情報検索提供装置 100 による検索結果を通じて知ることができる。海外企業と取引するとき、国内企業と比較して相手の素性が分からない場合が多いため、特に有意義である。さらに、企業の信用を調査する調査機関への取引先の信用調査依頼を省くことができるため、手間やコストを削減できる。

【0040】

本実施の形態に係る情報検索提供装置 100 によれば、ウェブページに連続してコンテンツを掲載しているほど信頼度を上昇させていく。そのため、ウェブサーバ 200 が昔から安定してウェブページを提供していればいるほど、信頼度が高く設定されようになる。ウェブページが、例えばインターネットショッピングサービスを提供するショッピングモールのページであり、そのウェブページに掲載されるコンテンツの信頼度が高ければ、ユーザはウェブページを提供する提供者が以前から安定して商売を続けていることを知ることができる。逆にその信頼度が低ければ、ウェブページの提供者は商売を開始して間もないか、ウェブページを転々とするいわゆる悪徳業者である可能性が高い。

【0041】

本実施の形態に係る情報検索提供装置 100 によれば、信頼のおけるコンテンツと信頼のおけないコンテンツとを信頼度をもとに選別してユーザに提示することで、ユーザは検索結果の中から信頼度の高いコンテンツのみを選択できる。これにより、ユーザは信頼度の高いコンテンツを選択することで、優良業者から提供されるサービスを利用できる。その結果、情報検索提供装置 100 はユーザ利便性の高い検索サービスを実現できる。

【0042】

履歴増加幅制御部 132 は、収集部 128 による評価対象のコンテンツの収集の可否の履歴に応じて当該評価対象のコンテンツの信頼度の増加幅を制御する。例えば、「2004 年 1 月 1 日」以降、収集部 128 により評価対象のコンテンツが収集され、その後、「2004 年 2 月 1 日」以降、収集部 128 により評価対象のコンテンツが収集されず、さらにその後、「2004 年 3 月 1 日」以降、収集部 128 により評価対象のコンテンツが収集されたとき、履歴増加幅制御部 132 は、「2004 年 3 月 1 日」における評価対象のコンテンツの信頼度の増加幅を、「2004 年 1 月 1 日」における評価対象のコンテンツの信頼度の増加幅よりも低く制御する。なお、履歴増加幅制御部 132 は、一定期間が経過すれば、低く制御したコンテンツの信頼度の増加幅を、低く制御しないときに想定されるもとの増加幅に戻してもよい。具体的には、「2004 年 3 月 1 日」での増加幅の減少分を、1 年後の「2005 年 3 月 1 日」に、現在の信頼度の増加幅に加えてもよい。

【0043】

これにより、昔から継続してコンテンツが収集可能であったときのウェブページの信頼度の増加幅は、途中、コンテンツが収集不可能となったときのウェブページ信頼度の増加幅よりも大きくなり、両者の増加幅を差別化することができる。その結果、コンテンツを継続して提供することの重要性をさらに高めることができる。

【0044】

図 4 は、検索情報保持部 180 に格納される検索情報ファイルのデータ構造の一例を示す。図 4 における検索情報ファイルは、IP アドレス欄 302 と、URL 欄 304 と、コンテンツ欄 306 と、ウェブページ提供開始日欄 308 と、信頼度欄 310 と、増加幅欄 312 とを含む。収集部 128 により収集されたコンテンツはコンテンツ欄 306 に、コンテンツ付随情報の一部の情報である URL は URL 欄 304 に格納される。

【0045】

IP アドレス欄 302 は、ネットワーク 12 に接続されるウェブサーバ 200 の IP アドレスを含む。本図の IP アドレス欄 302 に格納されている IP アドレスは IP v4 形式であるが別の例として IP v6 形式であってもよい。URL 欄 304 はネットワーク 1

2上におけるウェブページの所在を示す情報、いわゆるURLを備える。コンテンツ欄306には、収集部128により収集されたコンテンツ、本図では、HTML(HyperText Markup Language)言語で記述されたテキストデータを含む。ウェブページ提供開始日欄308はウェブサーバ200がウェブページの提供を開始した日を含む。信頼度欄310は、ウェブページに掲載されるコンテンツの信頼度を含む。増加幅欄312は信頼度の増加幅を含む。

【0046】

例えば、データ314について言えば、ウェブサーバ200のIPアドレスが「200.100.100.001」、ウェブサーバ200上のウェブページのURLが「http://~1/」、ウェブページの提供開始日は「2004年1月1日」、現在のコンテンツの信頼度は「600」で、信頼度の増加幅は「5」であることを示す。すなわち、URLが「http://~1/」であるウェブページに掲載されるコンテンツの信頼度は、コンテンツがそのウェブページ内に掲載される期間内において、所定のタイミングで、ここでは一日に一回「5」ずつ増加されることを示す。

【0047】

図5は、アクセス履歴保持部190に格納されたアクセス履歴ファイルのデータ構造の一例を示す。図4と同等の構成には同じ符号を与え適宜説明を略す。図5におけるデータ表は、IPアドレス欄302とアクセス履歴欄322を含む。アクセス履歴欄322は、アクセス部122によりウェブサーバ200にアクセスできたか否かを示す情報をアクセス履歴として保持する。具体的には、特定の日にフラグ「○」が格納されていれば、当該日はウェブページへのアクセスが可能であったことを示し、一方、フラグ「×」が格納されていれば、ウェブページへのアクセスが不可能であったことを示す。

【0048】

本図では便宜のため、「1月2日」から「5月1日」までの期間、「5月5日」から「5月30日」までの期間、および「6月3日」から「8月31日」までの期間についてのアクセス履歴の表示を省略しているが、それぞれの期間内におけるアクセス履歴は、それぞれの期間の開始日の前日におけるフラグの値をそのまま保持するものとする。なお、本図では「2004年1月1日」から現在の日付である「2004年9月1日」までのアクセス履歴を示しているが、もちろん、アクセス履歴欄322は「2003年12月31日」以前のアクセス履歴を備えてもよい。

【0049】

例えば、第1データ324について言えば、ウェブサーバ200のIPアドレスが「200.100.100.001」であり、「1月1日」から現在の日付である「9月1日」まで、アクセス部122によるウェブページへのアクセスが可能であったことを示している。すなわち、ウェブサーバ200はその期間中、継続的にウェブページを提供していることを示す。一方、第2データ326について言えば、「5月31日」までウェブページへのアクセスが不可能で、「6月1日」以降、ウェブページへのアクセスが可能であったことを示している。すなわち、IPアドレスが「200.100.100.002」であるウェブサーバ200上のウェブページが「6月1日」に提供開始されたことを示す。さらに、第3データ328によれば、IPアドレスが「200.100.100.003」であるウェブサーバ200は「5月3日」にウェブページの提供を開始したが、「5月4日」以降は、ウェブページを提供していない。なお、信頼度の増減制御の開始日は、例えば第2データ326について言えば、「6月1日」である。

【0050】

図6は、本実施の形態に係るコンテンツ収集処理の流れを示す。情報検索提供装置100は、アクセス履歴ファイルに記録されているすべてのIPアドレスについて、一日に一回、以下のコンテンツ収集処理を行う。アクセス部122は、アクセス履歴保持部190に格納されているアクセス履歴ファイルの中からコンテンツ収集対象のウェブサーバ200のIPアドレスを取得し(S10)、ウェブサーバ200にアクセスする。アクセス可否判断部124は、アクセス部122によるウェブサーバ200へのアクセスの可否を判

断し、アクセスが不可能であったと判断したとき（S 1 2 の N）、当該 I P アドレスでかつアクセス履歴欄 3 2 2 内の現在の日付の欄にフラグ「X」を設定する（S 1 4）。

【 0 0 5 1 】

アクセス可否判断部 1 2 4 は、アクセス部 1 2 2 によるアクセスが可能であったと判断したとき（S 1 2 の Y）、当該 I P アドレスでかつアクセス履歴欄 3 2 2 内の現在の日付の欄にフラグ「○」を設定する（S 1 8）。タイミング特定部 1 2 6 はアクセス履歴ファイルを参照し、現在の日付がウェブページの提供開始日に相当すると判断したとき（S 2 0 の Y）、ウェブページ提供開始日欄 3 0 8 に現在の日付を設定する。タイミング特定部 1 2 6 は現在の日付がウェブページの提供開始日に相当しないと判断したとき（S 2 0 の N）、ウェブページ提供開始日欄 3 0 8 に現在の日付を設定する動作をスキップする。

【 0 0 5 2 】

収集部 1 2 8 により当該 I P アドレスのウェブサーバ上のウェブページに掲載されるコンテンツの収集が可能であったとき（S 2 4 の Y）、収集部 1 2 8 は検索情報ファイル内のコンテンツ欄 3 0 6 にコンテンツを格納する（S 2 6）。このとき、コンテンツ付随情報も格納する。履歴増加幅制御部 1 3 2 はアクセス履歴ファイルを参照し、現在の日付の前日にアクセス部 1 2 2 によるアクセスが不可能であったと判断したとき（S 2 8 の Y）、増加幅を「1」減少させ（S 3 0）、信頼度制御部 1 3 0 は現在のコンテンツの信頼度を増加幅分だけ増加させる（S 3 2）。履歴増加幅制御部 1 3 2 は、アクセス履歴ファイルを参照し、現在の日付の前日にアクセス部 1 2 2 によるアクセスが不可能でなかったと判断したとき（S 2 8 の N）、信頼度制御部 1 3 0 は現在のコンテンツの信頼度を増加幅分だけ増加させる（S 3 2）。

【 0 0 5 3 】

収集部 1 2 8 により当該 I P アドレスのウェブページに掲載されるコンテンツの収集が不可能であったとき（S 2 4 の N）、信頼度制御部 1 3 0 は現在のコンテンツの信頼度を減少させる（S 3 4）。収集制御部 1 2 0 は、アクセス履歴ファイル中のすべての I P アドレスに対しコンテンツ収集処理が終了したか否かを判断し、コンテンツ収集処理が終了していない I P アドレスがあれば（S 1 6 の N）、その I P アドレスについて、コンテンツ収集処理を実施する。終了していれば、情報検索提供装置 1 0 0 はその日のコンテンツ収集処理を終了する（S 1 6 の Y）。

【 0 0 5 4 】

図 7 は、評価対象のコンテンツが継続的に収集可能である場合の評価対象のコンテンツの信頼度と時間との関係を示す。縦軸はコンテンツの信頼度、横軸は時間を示す。図中の日付 D 0 はウェブページの提供開始日であり、当該日以降の日を一日経過するごとの日付を日付 D 1 ～ D 4 と示す。本実施の形態では、一日経過するごとにコンテンツの信頼度が上下するため、実際の信頼度は離散値として表現されるが、当然ながら時間の経過とともに連続的に信頼度の変動するとしてもよく、本図ではその様子を示す。図示のごとく、収集部 1 2 8 により継続して評価対象のコンテンツが収集できる期間内、すなわち日付 D 1 以降の期間内において、評価対象のコンテンツの信頼度が徐々に増加している。

【 0 0 5 5 】

図 8 は、評価対象のコンテンツが途中、収集不可能となった場合のコンテンツの信頼度と時間との関係を示す。縦軸はコンテンツの信頼度、横軸は時間を示す。図 7 と同等のものには同じ符号を与え適宜説明を略す。本図では、日付 D 0 にウェブページの提供が開始され、当該日以降そのウェブページに掲載されるコンテンツの収集が可能であり、日付 D 2 以降コンテンツの収集が不可能になるが、日付 D 3 以降再度コンテンツの収集が可能になるときの信頼度の変化の様子を示す。コンテンツの収集可能な期間である、日付 D 0 から日付 D 2 までの期間および日付 D 3 以降の期間内において、信頼度制御部 1 3 0 によりコンテンツの信頼度は徐々に増加されているが、履歴増加幅制御部 1 3 2 により後者の期間内における信頼度の増加幅は前者の期間内における信頼度の増加幅よりも小さく制御されている。

【 0 0 5 6 】

実施の形態 2

実施の形態 1 では評価対象のコンテンツの収集が可能であったか否かに応じて当該コンテンツの信頼度を増減させたが、実施の形態 2 では収集の可否の結果に加え、評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無の結果を評価対象のコンテンツの信頼度の増減に反映させる。図 9 は、実施の形態 2 に係る情報検索提供装置 100 の構成を示す。図 3 と同等の構成には同じ符号を与え適宜説明を略す。実施の形態 2 に係る情報検索提供装置 100 には、図 3 に示した実施の形態 1 に係る情報検索提供装置 100 と異なる構成要素である関連性判断部 134 がさらに設けられる。

【0057】

関連性判断部 134 は、信頼度の高いコンテンツ内に、当該コンテンツと評価対象のコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて、評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断する。このとき、例えば、関連性判断部 134 は、信頼度の高いコンテンツ内に、当該関連情報、例えば信頼度の高いコンテンツから評価対象のコンテンツへのリンクが含まれているかどうか判断する。具体的には、関連性判断部 134 は、信頼度の高いコンテンツに含まれるタグ内の「〜」に、評価対象のコンテンツの URL が含まれているかどうか判断する。他の例で言えば、関連性判断部 134 は、信頼度の高いコンテンツ内の出店情報欄、紹介情報欄、提携情報欄や推薦情報欄内などに、評価対象のコンテンツからのリンク、評価対象のコンテンツを掲載するウェブページの URL やそのウェブページの情報提供者の名称があるか否かを判断する。このとき、それらの情報は関連性判断部 134 により検索情報ファイルから取得される。

【0058】

信頼度制御部 130 は、関連性判断部 134 により関連性があると判断されたときにおける評価対象のコンテンツの信頼度の増加幅を、関連性がないと判断されたときにおける評価対象のコンテンツの信頼度の増加幅よりも高く制御する。これにより、例えば、信頼度の高いウェブページ内に評価対象のコンテンツへのリンクがある場合とない場合とで、評価対象のコンテンツの信頼度を差別化できる。

【0059】

信頼度制御部 130 は、複数の信頼度の高いコンテンツにわたって、評価対象のコンテンツの URL と合致する URL があれば合致する信頼度の高いコンテンツの数だけ信頼度の増加幅を増加させてもよい。例えば、三つの信頼度の高いコンテンツ内に、評価対象のコンテンツの URL と合致する URL があれば、評価対象のコンテンツの現在の増加幅にさらに「3」を加算する。これにより、より多くの信頼度の高いコンテンツ内に、信頼度の高いコンテンツから評価対象のコンテンツへのリンクが張られれば張られるほど、信頼度の増加幅を大きくすることができる。

【0060】

また、信頼度制御部 130 は、複数の信頼度の高いコンテンツにわたって、当該コンテンツ内の出店情報欄、紹介情報欄、提携情報欄や推薦情報欄内に、評価対象のコンテンツからのリンク、評価対象のコンテンツを掲載するウェブページの URL やそのウェブページの情報提供者の名称が含まれていれば、含まれる信頼度の高いコンテンツの数だけ評価対象のコンテンツの信頼度の増加幅を増加させてもよい。例えば、含まれる信頼度の高いコンテンツの数が二つであれば、評価対象のコンテンツの現在の増加幅にさらに「2」を加算する。

【0061】

信頼度の高いコンテンツの情報提供者は、評価対象のコンテンツの情報提供者から見て第三者であるため、通常、評価対象のコンテンツの情報提供者はその信頼度の高いコンテンツの内容を勝手に変更することはできない。よって、信頼度の高いコンテンツ内に、信頼度の高いコンテンツから評価対象のコンテンツへのリンクや評価対象のコンテンツを紹介する等の情報が掲載されていれば、その事実は尊重されるべきである。上述のごとく、本実施の形態に係る信頼度制御部 130 は、上述のリンクや情報が掲載されている場合に

おける信頼度の増加幅を、それらが掲載されていない場合における信頼度の増加幅よりも高く制御することで、両者の差別化を図っている。

【 0 0 6 2 】

図 1 0 は、評価対象のコンテンツの信頼度と時間との関係を示す。縦軸はコンテンツの信頼度、横軸は時間を示す。図 7 と同等のものには同じ符号を与え適宜説明を略す。図中実線 1 1 は、信頼度の高いコンテンツ内に評価対象のコンテンツへのリンクが張られていない場合の評価対象のコンテンツの信頼度の変化の様子を示し、一方、実線 1 2 は、日付 D 0 に関連性判断部 1 3 4 により関連性があると判断され、当該日以降信頼度の高いコンテンツ内に評価対象のコンテンツへのリンクが張られている場合の評価対象のコンテンツの信頼度の変化の様子を示す。実線 1 2 の場合における信頼度の増加幅は、実線 1 2 の場合における信頼度の増加幅よりも大きいいため、実線 1 2 は実線 1 1 の信頼度を上回る。

【 0 0 6 3 】

実施の形態 3

実施の形態 1 では評価対象のコンテンツの収集可否に応じて当該コンテンツの信頼度を増減させたが、実施の形態 3 では当該コンテンツが更新されているか否かに応じて当該コンテンツの信頼度を増減させる。図 1 1 は、実施の形態 3 に係る情報検索提供装置 1 0 0 の構成を示す。図 3 と同等の構成には同じ符号を与え適宜説明を略す。実施の形態 3 に係る情報検索提供装置 1 0 0 には、図 3 に示した実施の形態 1 に係る情報検索提供装置 1 0 0 の履歴増加幅制御部 1 3 2 のかわりに、新たな構成要素である比較部 1 3 6 および更新判断部 1 3 8 がさらに設けられる。

【 0 0 6 4 】

比較部 1 3 6 は、収集部 1 2 8 により収集された評価対象のコンテンツの内容と、当該コンテンツを掲載するウェブページと同一のウェブページから収集される次の評価対象のコンテンツの内容とを比較する。比較の際、比較部 1 3 6 は、例えば、両者のコンテンツに含まれる HTML 言語構造、テキスト文字、画像やプログラムを参照する。比較部 1 3 6 は内部に、HTML 言語構造、テキスト文字、画像やプログラムの変位量を解析する図示しない変位量解析部を備えてもよい。比較部 1 3 6 は、ウェブページ固有の情報である、例えば、企業名、住所や電話番号について比較してもよい。前回収集されたコンテンツに記載されていた企業名、住所や電話番号が次に収集されたコンテンツにおいて理由無く変更されていたとき、信頼度制御部 1 3 0 によりコンテンツの信頼度は減少される。

【 0 0 6 5 】

更新判断部 1 3 8 は、比較部 1 3 6 による比較の結果に基づいて評価対象のコンテンツが更新されたか否かを判断する。このとき、更新判断部 1 3 8 は、前回収集したコンテンツの内容と次に収集されたコンテンツの内容とに含まれる、例えば文字が所定数以上、相違があったときに、評価対象のコンテンツが更新されたと判断する。なお、1 文字でも相違した場合に当該コンテンツが更新されたと判断してもよい。

【 0 0 6 6 】

別の例として HTML 言語構造を挙げれば、前回収集した評価対象のコンテンツ内の特定の場所に位置していたタグが別のタグに置き換わっていたり、新たなタグが追加挿入されたりしたとき、更新判断部 1 3 8 は、評価対象のコンテンツに更新があったものと判断してもよい。さらに別の例として画像を挙げれば、コンテンツ内の人物の写真が 1 0 年前と変わっていない場合、更新されていないものと判断してもよい。さらに別の例として、コンテンツ内の日記に記載する文章が途切れた場合や、掲示板の不適切な使用に対し何らかの対策を行っていない場合、更新判断部 1 3 8 は評価対象のコンテンツが更新されていないものと判断してもよい。

【 0 0 6 7 】

信頼度制御部 1 3 0 は、更新判断部 1 3 8 による判断の結果に応じて評価対象のコンテンツの信頼度を所定のタイミングで、ここでは一日に一回、増減させる。すなわち、更新判断部 1 3 8 により評価対象のコンテンツの更新があったと判断されたとき、評価対象のコンテンツが更新されていると、一日に一回、連続して判断される期間内において、当該

コンテンツの信頼度を徐々に増加させる。一方、更新判断部138により評価対象のコンテンツの更新がないと判断されたとき、評価対象のコンテンツが更新されていないと、一日に一回、連続して判断される期間内において、当該コンテンツの信頼度を徐々に減少させる。

【0068】

本実施の形態に係る情報検索提供装置100によれば、ウェブページに掲載されるコンテンツが定期的に更新されるほど信頼度が上がる可能性が高い。そのコンテンツの信頼度をユーザに提示することで、ユーザはそのコンテンツがどの程度の信頼度を有するかを簡単に知ることができる。例えば、ユーザは、信頼度の高い、インターネットによるショッピングサービスのウェブページを利用することで、安心してショッピングできる。一方、ウェブページ提供者側は自身のコンテンツの信頼度を高めようと、定期的な更新を積極的に行うことがありうる。これにより、ネットワーク上に存在するウェブページが全体的により更新される可能性が高くなり、ひいてはインターネット全体の活性化にもつながる。

【0069】

図12(a)は、評価対象のコンテンツが連続して更新されるとき信頼度と時間との関係を示し、図12(b)は、評価対象のコンテンツが途中更新されなくなったときの信頼度と時間との関係を示す。縦軸はコンテンツの信頼度、横軸は時間を示す。図7と同等のものには同じ符号を与え適宜説明を略す。図12(a)では、日付D0以降コンテンツが毎日更新されるため、図中の信頼度は徐々に増加している。一方、図12(b)では、日付D0～日付D2の期間内ではコンテンツが毎日更新されるが、日付D2以降コンテンツが更新されなくなり、当該日以降、信頼度は徐々に減少している。

【0070】

実施の形態4

実施の形態3では評価対象のコンテンツが更新されているか否かに応じて当該コンテンツの信頼度を増減させたが、実施の形態4では当該コンテンツの更新有無の結果に加え、当該コンテンツが本来更新を必要とするべきものであるか否かの分析結果を評価対象のコンテンツの信頼度を増減に反映させる。図13は、本実施の形態4に係る情報検索提供装置100の構成を示す。図11と同等の構成には同じ符号を与え適宜説明を略す。実施の形態4に係る情報検索提供装置100には、図11に示した実施の形態3に係る情報検索提供装置100に新たな構成要素である更新分析部140がさらに設けられる。

【0071】

更新分析部140は、収集部128により収集された評価対象のコンテンツの内容を参照し、当該コンテンツが本来更新を必要とするべきものであるか否かを分析する。具体的には、更新分析部140は評価対象のコンテンツを「本来更新の必要のあるグループ」と「本来更新の必要のないグループ」のいずれかのグループに分類する。「本来更新の必要のあるグループ」とは、情報を高い頻度で更新することで有意となるコンテンツ、例えば、ニュース速報、天気、日記を提供するコンテンツを含むグループをいう。一方、「本来更新の必要のないグループ」とは、情報を更新しないことで有意となるコンテンツ、例えば、政治家の主張、会社の基本的運営方針を提供するコンテンツを含むグループをいう。

【0072】

具体的には、更新分析部140はコンテンツの中に「ニュース」や「天気」など頻度の高い更新で有意になるキーワードが含まれていれば、そのコンテンツを「本来更新の必要のあるグループ」に分類する。一方、コンテンツの中に「主張」や「基本方針」など更新なしで有意となるキーワードが含まれていれば、そのコンテンツを「本来更新の必要のないグループ」に分類する。コンテンツをそれらグループに分類するためのキーワードは、あらかじめ情報検索提供装置100内に登録されていてもよい。

【0073】

信頼度制御部130は、更新判断部138による判断の結果および更新分析部140による分析の結果に応じて、評価対象のコンテンツの信頼度を増減させる。すなわち、「本来更新の必要のあるグループ」に分類された評価対象のコンテンツについては、信頼度制

信頼度制御部130は、更新判断部138によりコンテンツが更新されていると判断されたときに信頼度を増加させ、コンテンツが更新されていないと判断されたときに信頼度を減少させる。一方、「本来更新の必要のないグループ」に分類された評価対象のコンテンツについては、信頼度制御部130は、更新判断部138によりコンテンツが更新されていないと判断されたときに信頼度を増加させ、コンテンツが更新されていると判断されたときに信頼度を減少させる。

【0074】

本実施の形態によれば、コンテンツの更新有無に加え、コンテンツが本来更新を必要とするべきものであるか否かを信頼度の増減の判断材料にすることで、コンテンツが更新されなくてもそのコンテンツが本来更新すべきものではないものであれば信頼度を増加させることができ、当該コンテンツを救済できる。

【0075】

実施の形態5

実施の形態3では評価対象のコンテンツが更新されているか否かに応じて当該コンテンツの信頼度を増減させたが、実施の形態5では当該コンテンツが更新されているかどうかを調べるとき、当該コンテンツ内にリンク切れが生じているか否かを調べる。図14は、実施の形態5に係る情報検索提供装置100の構成を示す。図11と同等の構成には同じ符号を与え適宜説明を略す。実施の形態5に係る情報検索提供装置には図11に示した実施の形態3に係る情報検索提供装置100に新たな構成要素であるリンク切れ判断部141がさらに設けられる。

【0076】

リンク切れ判断部141は収集部128により収集された評価対象のコンテンツ内に記述されたリンク先のウェブページにアクセスし、当該ウェブページに何らかのコンテンツが掲載されているか否かを判断する。具体的には、リンク切れ判断部141は、評価対象のコンテンツに含まれるタグ内の「～」を参照して他のウェブサーバのウェブページにアクセスする。上記のタグが複数あれば複数のウェブページにアクセスする。このとき、例えば「移動しました」や「FILE NOT FOUND」などの文字を発見したとき、リンク切れが生じていると判断する。一方、ウェブページに何らかのコンテンツ、例えば上記の文字以外を含むコンテンツが掲載されていれば、リンク切れが生じていないと判断する。

【0077】

もちろん、リンク切れ判断部141は、同一ウェブサーバ上の同一ウェブページへのリンク先や同一ウェブサーバ上の異なるウェブページへのリンク先にアクセスしてもよい。このとき、リンク切れ判断部141は、例えば、リンク先に画像や映像データが存在しないとき、リンク切れが生じていると判断する。

【0078】

信頼度制御部130はリンク切れ判断部141による判断の結果に応じて、評価対象のコンテンツの信頼度の増減を制御する。具体的には、信頼度制御部130はリンク切れが生じていると判断されれば、リンク切れが生じている期間内において、評価対象のコンテンツの信頼度を一定に保つ。なお、この場合、信頼度を減少させてもよい。一方、評価対象のコンテンツが更新されリンク切れが解消すればコンテンツの信頼度を増加させる。

【0079】

本実施の形態によれば、ウェブページ内にリンク切れがあるときの信頼度をウェブページ内にリンク切れがないときの信頼度よりも低くすることで、両者を差別化できる。ウェブページとウェブページとをつなぐリンクは、インターネットが発展した理由の重要な要素の一つであるため、リンク切れが有るか否かをコンテンツの信頼度を評価する際の指標にするのは有意義である。リンク切れに基づいた信頼度をユーザに提示することで、ユーザはリンク切れのない信頼度の高いコンテンツのみを利用できる。一方、ウェブページ提供者側はコンテンツの信頼度を高めようと、リンク切れの修正を積極的に行うことがありうる。これにより、ネットワーク上に存在するデッドリンクが全体的に見てより修正され

る可能性が高くなり、ひいてはウェブページ全体の品質の向上にもつながる。

【0080】

図15は、評価対象のコンテンツに、途中リンク切れが生じた場合のコンテンツの信頼度と時間との関係を示す。縦軸はコンテンツの信頼度、横軸は時間を示す。図7と同等のものには同じ符号を与え適宜説明を略す。本図では、日付D0にウェブページの提供が開始され、当該日以降そのウェブページに掲載されるコンテンツが一日に一回更新され、日付D2以降コンテンツにリンク切れが生じるが、日付D3に、コンテンツの更新によりリンク切れが修正されたときの信頼度の変化の様子を示す。図示のごとく、リンク切れが生じている期間である、日付D2から日付D3までの期間内において、信頼度制御部130によりコンテンツの信頼度が一定に保持されている。

【0081】

実施の形態6

実施の形態1ではコンテンツの収集が可能であったか否かに応じて評価対象のコンテンツの信頼度を増減させたが、実施の形態6では評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素に応じてコンテンツの信頼度を増減させる。図16は、実施の形態6に係る情報検索提供装置100の構成を示す。図3と同等の構成には同じ符号を与え適宜説明を略す。実施の形態6に係る情報検索提供装置100には、図3に示した実施の形態1に係る情報検索提供装置100の履歴増加幅制御部132のかわりに、新たな構成要素である関連コンテンツ取得部142、要素抽出部144および分類部146がさらに設けられる。

【0082】

関連コンテンツ取得部142は評価対象のコンテンツに関連するコンテンツを取得する。例えば、検索部114が、評価対象のコンテンツに含まれる検索条件文、例えば「ABC会社」文を指定して検索を行い、関連コンテンツ取得部142は、その検索条件文に合致したコンテンツを関連するコンテンツ（以下、「関連コンテンツ」という）としてもよい。このとき、検索部114は検索情報保持部180から関連コンテンツを取得してもよいし、他の検索エンジンやウェブページを利用して関連コンテンツを取得してもよい。検索条件文は、端末装置50を介してユーザから指定されてもよいし、情報検索提供装置100内部の図示しない検索条件文生成部により辞書を用いて定期的に生成されてもよい。なお、関連コンテンツは評価対象のコンテンツ自身を含んでもよい。

【0083】

要素抽出部144は取得されたコンテンツの中から評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出する。例えば、信頼度の増加に貢献する要素とは「新商品開発」、「特許取得」、「新規展開」、「増収増益」、などのキーワードを示し、信頼度の減少に貢献する要素とは「裁判」、「負債」、「逮捕」、「減益」などのキーワードを示す。信頼度の増加あるいは減少に貢献するキーワードのすべては、あらかじめ図示しない貢献保持部に保持されていてもよい。このとき、要素抽出部144は、その貢献保持部からキーワードを取得し、関連コンテンツ取得部142により取得されたコンテンツに対して、これらのキーワードを指定して検索し合致した情報を取得する。

【0084】

分類部146は要素抽出部144により抽出された要素それぞれを、評価対象のコンテンツの信頼度の増加に貢献する第1グループ、あるいは評価対象のコンテンツの信頼度の減少に貢献する第2グループに分類する。

【0085】

信頼度制御部130は要素抽出部144により抽出された要素を参照して評価対象のコンテンツの信頼度を増減させる。具体的には、信頼度制御部130は、分類部146により分類された第1グループ内の要素数が第2グループ内の要素数よりも多いとき、評価対象のコンテンツの信頼度を増加させる。一方、信頼度制御部130は、第2グループ内の要素数が第1グループ内の要素数よりも多いとき、評価対象のコンテンツの信頼度を減少させる。

【0086】

本実施の形態によれば、情報検索提供装置100は、評価対象のコンテンツが他のウェブページやコンテンツからどのように評価されているかに応じて評価対象のコンテンツの信頼度を増減できる。このとき、評価対象のコンテンツの信頼度の上昇に貢献する要素だけでなく、評価対象のコンテンツの信頼度の下降に貢献する要素をも考慮することで、信頼度の評価の精度が高まる。

【0087】

実施の形態7

実施の形態1～6では評価対象のコンテンツの信頼度を単独で増減させたが、実施の形態7ではお互いに関連するコンテンツの信頼度を同時に制御する。図17は、実施の形態7に係る情報検索提供装置100の構成を示す。図3と同等の構成には同じ符号を与え適宜説明を略す。実施の形態7に係る情報検索提供装置100には、図3に示した実施の形態1に係る情報検索提供装置100の履歴増加幅制御部132のかわりに、新たな構成要素であるコンテンツ取得部147、自コンテンツ特定部148および整合性判断部150がさらに設けられる。

【0088】

コンテンツ取得部147は、判断対象項目を含むコンテンツを取得する。判断対象項目とは、整合性の判断対象のデータであり、例えば住所、電話番号やファクシミリ番号、他に採用関連情報や財務関連情報などであってもよい。採用関連情報とは、例えば、来年の採用予定人数や今年の採用実績人数、財務関連情報とは、売上高や経常利益額などを指す。なお、判断対象項目のデータ形式は問わず、テキストデータ、画像データ、音声データや映像データであってもよい。自コンテンツ特定部148は、取得されたコンテンツのうち判断対象項目の情報発生源となる自コンテンツを特定する。以下、自コンテンツを特定する際の判断材料の一例を示す。もちろん、判断材料は以下の組み合わせで構成されてもよい。

- (1) コンテンツのURLを参照して特定する。
- (2) コンテンツの内容を参照して特定する。
- (3) 他のウェブページ、検索エンジンやデータベースから得られる検索結果を参照して特定する。
- (4) コンテンツに張られているリンクの総数で特定する。
- (5) コンテンツの提供開始日で特定する。

【0089】

(1)の場合、自コンテンツ特定部148は、取得されたコンテンツのURLを参照し、例えば、判断対象項目が特定の企業の電話番号であり、URL内にその企業のドメイン名「abc」を含んでいるか否かを判断し、それを含んでいればそのコンテンツを自コンテンツであると判断する。このとき、アクセス部122はドメイン名を管理する第三者機関のデータベースにアクセスして、その企業のドメイン名を取得してもよい。

【0090】

(2)の場合、自コンテンツ特定部148は取得されたコンテンツの内容を参照し、例えば、判断対象項目が特定の企業の電話番号であり、コンテンツ内のタイトル部分、例えば、タグ<TITLE>～</TITLE>に囲まれた領域に企業の名称があれば、自コンテンツであると判断する。

【0091】

(3)の場合、自コンテンツ特定部148は、判断対象項目を検索条件文と指定して他のウェブページ、検索エンジンやデータベースから得られる検索結果のうち最上位に位置するコンテンツを自コンテンツと判断する。(4)の場合、自コンテンツ特定部148は、コンテンツ内のタグを参照し、他のコンテンツから張られるリンクの数が一番多いコンテンツを自コンテンツであると判断する。

【0092】

(5)の場合、自コンテンツ特定部148は取得されたコンテンツに関するウェブペー

ジの提供開始日を参照し、最先の提供開始日であるコンテンツを自コンテンツであると判断する。

【 0 0 9 3 】

整合性判断部 1 5 0 は、自コンテンツ特定部 1 4 8 により特定された自コンテンツの内容と取得されたコンテンツのうち自コンテンツ以外の他コンテンツの内容とを比較し、判断対象項目の整合性の有無を判断する。

【 0 0 9 4 】

信頼度制御部 1 3 0 は、整合性判断部 1 5 0 による判断の結果に応じて自コンテンツの信頼度と他コンテンツの信頼度を増減させる。すなわち、信頼度制御部 1 3 0 は、自コンテンツの信頼度および他コンテンツの信頼度がともに高いときであって、整合性判断部 1 5 0 により判断対象項目について整合性有りとは判断されたとき、信頼度制御部 1 3 0 は自コンテンツの信頼度および他コンテンツの信頼度をともに増加させる。自コンテンツの信頼度および他コンテンツの信頼度がともに高いときであって、整合性無しとは判断されたとき、信頼度制御部 1 3 0 は両方のコンテンツの信頼度を下げる。このとき、情報検索提供装置 1 0 0 は図示しない通知部を備え、それぞれのコンテンツを掲載するウェブページの情報提供者に判断対象項目の是正の指示を通知してもよい。

【 0 0 9 5 】

自コンテンツの信頼度が高く他コンテンツの信頼度が低いときであって、整合性有りとは判断されたとき、信頼度制御部 1 3 0 は他コンテンツの信頼度を上げるが、自コンテンツの信頼度については増減させない。さらに、自コンテンツの信頼度が高く他コンテンツの信頼度が低いときであって、整合性無しとは判断されたとき、信頼度制御部 1 3 0 は他コンテンツの信頼度を下げるが、自コンテンツの信頼度については増減させない。すなわち、信頼度が低い他コンテンツはそもそも信頼のおけない内容であるため、信頼度制御部 1 3 0 はそれに左右されることがないように自コンテンツの信頼度を制御するのである。

【 0 0 9 6 】

図 1 8 は、自コンテンツと他コンテンツとの判断対象項目に関する関係を模式的に示す。自コンテンツ R 1 は判断対象項目である「情報 X＝鹿児島県」、他コンテンツ R 2 は「情報 X＝鹿児島県」、他コンテンツ R 3 は「情報 X＝沖縄県」を含む。整合性判断部 1 5 0 は、自コンテンツ R 1 と他コンテンツ R 2 とを比較し、あるいは自コンテンツ R 1 と他コンテンツ R 3 とを比較し、整合性の有無を判断する。本図の場合、整合性判断部 1 5 0 により自コンテンツ R 1 と他者コンテンツ R 2 については整合性有りとは判断され、自コンテンツ R 1 と他コンテンツ R 3 については整合性無しとは判断される。

【 0 0 9 7 】

他コンテンツは公のあるいは公共性の高い事業者により提供されるコンテンツであってもよい。公のあるいは公共性の高い事業者とは、例えば国、地方公共団体や第三者機関により運営される事業者を指す。それらの事業者により提供されるコンテンツの信頼度は一般的に高いため、整合性判断部 1 5 0 は、判断対象項目、例えば電話番号や住所について、自コンテンツと他コンテンツとの整合性があると判断すれば、自コンテンツの信頼度を増加させることができる。

【 0 0 9 8 】

一般的にコンテンツの中には虚偽や誤りの内容を含んだものもある。そのため、情報提供者が発信する内容について他のコンテンツやウェブページに掲載されている内容との整合性の有無を判断することは有意義である。本実施の形態に係る情報検索提供装置 1 0 0 により、整合性の有無に基づいたコンテンツの信頼度を通知されることで、ユーザはそのコンテンツが虚偽や誤りの内容を含む可能性の少ない、信頼度の高いコンテンツのみを利用できる。一方、ウェブページ提供者側はコンテンツの信頼度を高めようと、虚偽や誤りの修正を積極的に行うことがありうる。これにより、コンテンツ全体の虚偽や誤りが全体的に見てより修正される可能性が高くなり、これによりウェブページ全体の品質の向上にもつながる。

【 0 0 9 9 】

実施の形態 8

実施の形態 8 に係る情報検索提供装置 100 は、実施の形態 1、実施の形態 3、実施の形態 5、実施の形態 6 および実施の形態 7 に係る機能を実現する構成要素を備える。このとき、それぞれの実施の形態に係る構成要素で実現される信頼度は、それぞれ異なる信頼度として検索情報ファイル内の信頼度欄 310 に管理される。なお、実施の形態 8 に係る情報検索提供装置 100 は、実施の形態 1 に係る機能を実現する構成要素にかえて実施の形態 2 に係る機能を実現する構成要素、実施の形態 3 に機能を実現する構成要素にかえて実施の形態 4 に係る機能を実現する構成要素を備えてもよい。

【0100】

図 19 は、情報検索提供装置 100 が提供する検索初期画面の一例を示す。検索初期画面 400 は、検索条件文を受け付ける第 1 入力ボックス 402 と第 2 入力ボックス 404 と第 3 入力ボックス 406、開始ボタン 408、およびキャンセルボタン 410 などで構成される。第 1 入力ボックス 402 には、会社名や代表者名や商品名に関する検索条件文が入力され、第 2 入力ボックス 404 には URL や電話番号やファクシミリ番号や住所が入力され、第 3 入力ボックス 406 には、いわゆるフリー検索を行う際の検索条件文がユーザにより入力される。少なくとも一つのボックスに検索条件文が入力された後、開始ボタン 408 が押下されたとき、入力された検索条件文が送信部 52 により情報検索提供装置 100 に送信され、検索処理が行われる。送信中にキャンセルボタン 410 が押下されれば、情報検索提供装置 100 は検索処理を停止する。

【0101】

図 20 は、情報検索提供装置 100 が提供する検索結果画面 401 の一例を示す。検索結果画面 401 には、検索条件文に合致したコンテンツと当該コンテンツに関連する各種情報、例えば信頼度が表示される。検索結果画面 401 は、コンテンツを掲載するウェブページの情報提供者が表示される領域 412、ウェブページの提供開始日が表示される領域 414、信頼度と時間との関係が表形式で表示される信頼度表 420、表示されたコンテンツに対する評価内容が記載されている他コンテンツの一覧を示す一覧領域 430、および特定の項目における信頼度をアルファベット形式で表示される信頼度ランク表 440 等で構成される。

【0102】

信頼度表 420 は 5 つの第 1 ～第 5 信頼度表 420a ～420e を備える。ここで、第 1 信頼度表 420a は実施の形態 1 あるいは実施の形態 2 に係るコンテンツの連続性に関する信頼度表、第 2 信頼度表 420b は実施の形態 3 あるいは実施の形態 4 に係るコンテンツの更新状況に関する信頼度表、第 3 信頼度表 420c は実施の形態 5 に係るコンテンツのリンク先状況に関する信頼度表、第 4 信頼度表 420d は実施の形態 6 に係るコンテンツの他コンテンツからの評価に関する信頼度表、第 5 信頼度表 420e はそれら第 1 ～第 4 信頼度表 420a ～420d の結果を総合化した信頼度表である。5 つの第 1 ～第 5 信頼度表 420a ～420e 中の点 A は現在のコンテンツの信頼度を示す。

【0103】

一覧領域 430 は、検索されたコンテンツに対し良好と評価している他ウェブページの一覧が格納される第 1 一覧領域 422 と検索されたコンテンツに対し良好でないと評価している他ウェブページの一覧が格納される第 2 一覧領域 424 を備える。具体的には、実施の形態 6 に係る関連コンテンツ取得部 142 により取得されたコンテンツのうち、評価対象のコンテンツの信頼度の増減に貢献するコンテンツが表示されている。図示のごとく、第 1 一覧領域 422 には検索されたコンテンツの信頼度を増加させるキーワードを含むコンテンツを掲載するウェブページの一覧が格納され、第 2 一覧領域 424 には検索されたコンテンツの信頼度を減少させるキーワードを含むコンテンツを掲載するウェブページの一覧が格納されている。

【0104】

信頼度ランク表 440 は、項目欄 442 およびランク欄 444 を含む。この表は、実施の形態 7 に係る構成要素による整合性判断の結果として制御される信頼度が表示されたも

のである。この信頼度ランク表 4 4 0 は、前述の判断対象項目である住所、電話番号、ファクシミリ番号や採用関連情報などを含む。ランク欄 4 4 4 に含まれる情報は、コンテンツの信頼度をその大きさに応じていくつかのカテゴリに分類し、それをアルファベット形式で示したものである。例えば、ランク「A」は信頼度が高い状態を示し、以下順番に信頼度が低くなるにつれ、アルファベット順に「B」、「C」、「D」のように示す。

【0105】

本実施の形態によれば、複数の観点、例えばコンテンツの連続性やコンテンツの更新状況の観点におけるコンテンツの信頼度をユーザに提示することで、ユーザは様々な観点からのコンテンツの信頼度を知ることができる。特に総合評価の信頼度は、複数の観点からの信頼度をまとめたものであり、バランス面において優れており、ユーザはコンテンツに関するより信頼性の高い評価結果を知ることができる。

【0106】

以上、本発明を実施の形態をもとに説明した。実施の形態は例示であり、それらの各構成要素や各処理プロセスの組み合わせにいろいろな変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。以下、変形例を挙げる。

【0107】

変形例 1

実施の形態に係る情報検索提供装置 1 0 0 は一つの装置として存在されていたが、変形例の構成として、情報検索提供装置 1 0 0 を 2 つのサーバの協働によって実現してもよい。一例として、図 3 の検索制御部 1 1 0 を検索用サーバ側に設け、それら以外の機能ブロックを収集用サーバ側に置いてよい。当然ながら、図 3 の構成をこれ以外の部分で分割して検索用サーバ側と収集用サーバ側に分配することができる。どのような分配がよいかについては、サーバ管理者の管理の利便性、2 つのサーバの負荷バランス、両者の通信量などの要素を考慮して定めることができる。

【0108】

変形例 2

実施の形態 8 に係る情報検索提供装置 1 0 0 は、実施の形態 1、実施の形態 3、実施の形態 5、実施の形態 6 および実施の形態 7 に係る情報検索提供装置 1 0 0 が有する構成要素を備えたが、別の変形例の構成として、その組み合わせはこれに限定されない。例えば、実施の形態 1 に係る情報検索提供装置 1 0 0 の構成要素と実施の形態 2 に係る情報検索提供装置 1 0 0 の構成要素を含む装置を新たな情報検索提供装置 1 0 0 としてもよい。

【0109】

変形例 3

実施の形態 3 では評価対象のコンテンツが更新されているか否かに応じて当該コンテンツの信頼度を増減させたが、さらに別の変形例として、情報検索提供装置 1 0 0 は当該コンテンツの内容を分析する分析部を備え、分析部による分析の結果、評価対象のコンテンツの信頼度を上昇させる項目を含んでいる場合は、信頼度を評価対象のコンテンツの信頼度を増加させてもよい。以下、当該項目の一例を示す。なお、この分析部は内部に時間情報を取得するためのタイマやプログラムのソースを解析するソース解析部を備えるものとする。

(ア) コンテンツに SSL (Secure Sockets Layer) 対策などのセキュリティ対策を行っているとき

(イ) コンテンツがクレジット決済を扱っているとき

(ウ) コンテンツが有料ドメイン名を使用しているとき

(エ) コンテンツが悪質なプログラムを用いていないとき

(オ) コンテンツの利用者およびアクセス数が多いとき

(カ) コンテンツへアクセスした際の応答時間が短いとき

(キ) コンテンツ内の電子メールアドレスに当該コンテンツのウェブページのドメイン名が含まれているとき

(ク) 性的な内容や暴力的な内容など公序良俗に反する内容を含んでいないとき

【0110】

変形例4

実施の形態では、検索情報保持部180に保持されているコンテンツすべてを検索結果の提示対象としていたが、さらに別の変形例として、情報検索提供装置100は、検索結果のうち信頼度の高いコンテンツしかユーザに提示させない提示制限部を備えてもよい。インターネット上には悪質なウェブページもあり、信頼度が低いコンテンツをそもそもユーザに提示させないことによって、ユーザ全体における信頼度の低いウェブページへのアクセスを軽減できる。

【0111】

変形例5

実施の形態7では、判断対象項目、例えば採用関連情報の整合性の有無に応じて、お互いに関連する自コンテンツと他コンテンツの信頼度を同時に制御したが、さらに別の変形例として、コンテンツ取得部147により取得されたコンテンツのうち、判断対象項目を掲載しているコンテンツが多いほど自コンテンツの信頼度を上げてても良い。例えば、判断対象項目が採用関連情報であり、「新卒募集」であったとき、信頼度制御部130は、その「新卒募集」を含むコンテンツが多いほど、自コンテンツの信頼度を増加させる。

【0112】

変形例6

さらに別の変形例として、実施の形態1～8に係る情報検索提供装置100に将来のコンテンツの信頼度を予測する予測部をさらに設けてもよい。予測部は、評価対象のコンテンツに含まれるコンテンツの存続期間の推測に供する要素に基づいて、評価対象のコンテンツの信頼度を予測する。コンテンツの存続期間の推測に貢献する要素とは、例えば、コンテンツの情報提供者が個人であれば、当該個人の年齢に相当し、一方、企業であれば、会社の成長率に相当する。前者の例でいえば、個人の年齢が「30歳」であれば、通常の人々の寿命である「80歳」まで、すなわち現在から50年後のコンテンツの信頼度を予測する。予測の際、現在の信頼度の増加幅を維持するという条件のもとで、50年後のコンテンツの信頼度を予測してもよい。信頼度提示部118はその予測された信頼度をユーザに提示する。これにより、ユーザは現在、過去の信頼度の変化の様子を知るだけでなく、将来のコンテンツの信頼度を知ることによって、信頼度がさらに増加していれば、より安心してコンテンツを利用することができる。

【0113】

変形例7

実施の形態1では、収集部128によるコンテンツの収集の可否に応じて信頼度の増減を制御したが、さらに別の変形例として、収集部128によりコンテンツが収集されないとき、情報検索提供装置100に新たに設けられた確認部が、確認対象項目についてコンテンツの収集の否に関する確認を行ってもよい。以下、確認対象の項目の一例を示す。サーバ移転情報やサーバ故障情報に含まれるキーワードは、図示しない確認対象項目保持部にあらかじめ保持されている。

(A) コンテンツ内に「移転しました」などのサーバ移転情報があるか否か

(B) コンテンツ内に「ウェブサーバ故障中」などのサーバ故障情報があるか否か

【0114】

それらサーバ移転情報やサーバ故障情報が評価対象のコンテンツ内に含まれているとき、評価対象のコンテンツが収集されなくても、信頼度制御部130は、評価対象のコンテンツの信頼度を減少させないように制御してもよい。また、このとき、アクセス可否判断部124は、その情報がコンテンツ内に掲載されている期間内において、評価対象のコンテンツのアクセス履歴欄322内にフラグ「○」を設定する。また、例えば、評価対象のコンテンツの収集の否が10日間続いたとき、確認部はドメイン名を管理する第三者機関のデータベースにアクセスして、評価対象のコンテンツを掲載するウェブサーバのドメイン名をキーにして、当該ウェブサーバを提供する情報提供者の状態を把握してもよい。これ

により、情報提供者の責めに期する理由がなく、情報提供者がコンテンツを提供できなかったとき、当該コンテンツの信頼度の減少をさせないことで、当該コンテンツを救済できる。

【0115】

変形例 8

実施の形態 1～8 では、情報検索提供装置 100 は検索結果としてコンテンツとコンテンツの信頼度の一覧をユーザに示したが、さらに別の変形例として、ユーザが、その情報を用いて、例えば実際の店舗に赴いたり、あるいはインターネットショッピングで売買をしたりしたとき、情報検索提供装置 100 はユーザからのそのコンテンツに対する信頼度に関する評価値を取得し、現在の当該コンテンツの信頼度に反映させる。本変形例では、実施の形態 1～8 に係る情報検索提供装置 100 に、新たに評価値取得部が設けられる。

【0116】

評価値取得部はユーザからのコンテンツに対する信頼度に関する評価値を取得する。その評価値取得部を介して、ユーザからコンテンツの信頼度に関する、低い方から高い方へ「-5」～「5」の 10 段階評価で表現された評価値を取得し、信頼度制御部 130 はその評価値を現在のコンテンツの信頼度に加算する。これにより、コンテンツを利用した際のユーザからの当該コンテンツに対する評価を取り込むことができ、よりの確な信頼度を提示できる情報検索サービスを実現できる。

【0117】

変形例 9

実施の形態 2 では、信頼度の高いコンテンツ内に、当該コンテンツと評価対象のコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて関連性の有無を判断したが、本変形例では、評価対象のコンテンツ内に、評価対象のコンテンツと信頼度の高いコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて、関連性の有無を判断する。例えば、関連性判断部 134 は、評価対象のコンテンツ内に、当該関連情報、例えば、評価対象のコンテンツから信頼度の高いコンテンツへのリンクが含まれているかどうか判断する。具体的には、関連性判断部 134 は、評価対象のコンテンツに含まれるタグ内の「～」に、信頼度の高いコンテンツを掲載するウェブページの URL が含まれているかどうか判断する。このとき、信頼度の高いコンテンツを掲載するウェブページの URL は、関連性判断部 134 により検索情報保持部 180 に保持された検索情報ファイルから取得される。

【0118】

信頼度制御部 130 は、関連性判断部 134 により関連性があると判断されたときにおける評価対象のコンテンツの信頼度の増加幅を、関連性がないと判断されたときにおける評価対象のコンテンツの信頼度の増加幅よりも高く制御する。これにより、例えば、評価対象のコンテンツ内に信頼度の高いウェブページへのリンクがある場合と、信頼度の高いウェブページへのリンクがない場合とを差別化できる。信頼度制御部 130 は、評価対象のコンテンツ内に、信頼度の高い複数のコンテンツの URL と合致するものがあれば合致する数だけ信頼度の増加幅を増加させてもよい。例えば、合致するものが三つあれば、評価対象のコンテンツの現在の増加幅にさらに「3」を加算する。これにより評価対象のコンテンツに信頼度の高いコンテンツへのリンクが張られれば張られるほど、信頼度の増加幅を大きくすることができる。

【0119】

変形例 10

実施の形態 2 および変形例 9 では、評価対象のコンテンツと信頼度の高いコンテンツとの関連性の有無を判断したが、さらに別の変形例として、評価対象のコンテンツと信頼度の低いコンテンツとの関連性の有無を判断する。本変形例における関連性判断部 134 による関連性の有無の判断方法については、実施の形態 2 および変形例 9 における関連性の有無の判断方法と同様である。すなわち、関連性判断部 134 は、信頼度の低いコンテンツ内に、当該コンテンツと評価対象のコンテンツとの間に関連性があることを示す関連情

報が含まれるか否かに基づいて、評価対象のコンテンツと信頼度の低いコンテンツとの関連性の有無を判断する。さらに、関連性判断部134は、評価対象のコンテンツ内に、当該コンテンツと信頼度の低いコンテンツとの間に関連性があることを示す関連情報が含まれるか否かに基づいて、評価対象のコンテンツと信頼度の低いコンテンツとの関連性の有無を判断してもよい。

【0120】

前者の判断の場合、信頼度制御部130は、関連性判断部134による関連性の有無の判断の結果を、評価対象のコンテンツの信頼度の増加幅の増減に反映させない。信頼度の低いコンテンツは信頼のおけない情報を有する可能性が高く、リンクなどの関連情報の存在自体を無視した方が有意義であるためである。後者の判断の場合、信頼度制御部130は、関連性判断部134により関連性があると判断されたときにおける評価対象のコンテンツの信頼度の増加幅を、関連性がないと判断されたときにおける評価対象のコンテンツの信頼度の増加幅よりも低く制御する。これにより、例えば、評価対象のコンテンツ内に、性的な内容や暴力的な内容など公序良俗に反する内容を含んだ、信頼度の低いコンテンツへのリンクが有る場合と、信頼度の低いウェブページへのリンクがない場合とを差別化できる。

【0121】

変形例11

実施の形態6では単一の観点から評価対象のコンテンツの信頼度を増減させたが、本変形例では、複数の観点から評価対象のコンテンツの信頼度を増減させる。このとき、要素抽出部144は、信頼度の評価の観点ごとに、前記評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出する。具体的には、信頼度の評価の観点が、例えば「採用」であれば、要素抽出部144は、関連コンテンツ取得部142により取得されたコンテンツの中から、「採用」に関連する、評価対象のコンテンツの信頼度の増加あるいは減少に貢献する要素を抽出する。

【0122】

信頼度の評価の観点、例えば「採用」に関するキーワードは、あらかじめ図示しない観点保持部に保持されていてもよい。「採用」に関するキーワードとは例えば、「採用実績人数」や「採用予定人数」などが挙げられる。このとき、要素抽出部144は、関連コンテンツ取得部142により取得されたコンテンツに対して、観点保持部の保持されたキーワードおよび前述の貢献保持部に保持されたキーワードを指定して検索し合致した情報を取得する。具体的には、要素抽出部144は、「採用実績人数」と「増加」をAND条件として検索すれば、「採用」の観点から信頼度の増加に貢献する要素を抽出できる。

【0123】

信頼度制御部130は、信頼度の評価の観点ごとに抽出された要素を参照して、評価対象のコンテンツの信頼度を増減せしめる。例えば、評価対象のコンテンツが「ABC会社」により提供されるものであり、取得されたコンテンツの中に「ABC会社の採用人数は増加した」との記載があれば、「採用」という観点から「増加」という要素が要素抽出部144により抽出され、その要素は分類部146により前述の第1グループに分類される。このとき、信頼度制御部130は、「採用」という信頼度の評価の観点から評価対象のコンテンツの信頼度を増加せしめる。なお、信頼度制御部130は、観点ごとに信頼度を設定し、要素抽出部144による信頼度の評価の観点ごとに抽出された結果に基づいて、観点ごとにそれぞれの信頼度を増減せしめてもよい。このとき、信頼度制御部130は、観点ごとに得られたそれぞれの信頼度を合計して、一つの評価対象のコンテンツの信頼度としてもよい。本変形例によれば、評価対象のコンテンツの信頼度を様々な観点から評価でき、信頼度の精度が高まる。

【図面の簡単な説明】

【0124】

【図1】実施の形態1に係る情報検索提供システムの構成を示す図である。

【図2】実施の形態1に係る端末装置の構成を示す図である。

【図 3】実施の形態 1 に係る情報検索提供装置の構成を示す図である。

【図 4】検索情報保持部に格納される検索情報ファイルのデータ構造の一例を示す図である。

【図 5】アクセス履歴保持部に格納されたアクセス履歴ファイルのデータ構造の一例を示す図である。

【図 6】実施の形態 1 に係るコンテンツ収集処理の流れを示す図である。

【図 7】継続的に収集可能である場合の評価対象のコンテンツの信頼度と時間との関係を示す図である。

【図 8】評価対象のコンテンツが途中、収集不可能となった場合のコンテンツの信頼度と時間との関係を示す図である。

【図 9】実施の形態 2 に係る情報検索提供装置の構成を示す図である。

【図 10】評価対象のコンテンツの信頼度と時間との関係を示す図である。

【図 11】実施の形態 3 に係る情報検索提供装置の構成を示す図である。

【図 12】図 12 (a) は、評価対象のコンテンツが連続して更新されるとき信頼度と時間との関係を示し、図 12 (b) は、評価対象のコンテンツが途中更新されなくなったときの信頼度と時間との関係を示す図である。

【図 13】実施の形態 4 に係る情報検索提供装置の構成を示す図である。

【図 14】実施の形態 5 に係る情報検索提供装置の構成を示す図である。

【図 15】評価対象のコンテンツに途中、リンク切れが生じた場合のコンテンツの信頼度と時間との関係を示す図である。

【図 16】実施の形態 6 に係る情報検索提供装置の構成を示す図である。

【図 17】実施の形態 7 に係る情報検索提供装置の構成を示す図である。

【図 18】自己コンテンツと他者コンテンツとの信頼度の相関関係を模式的に示す図である。

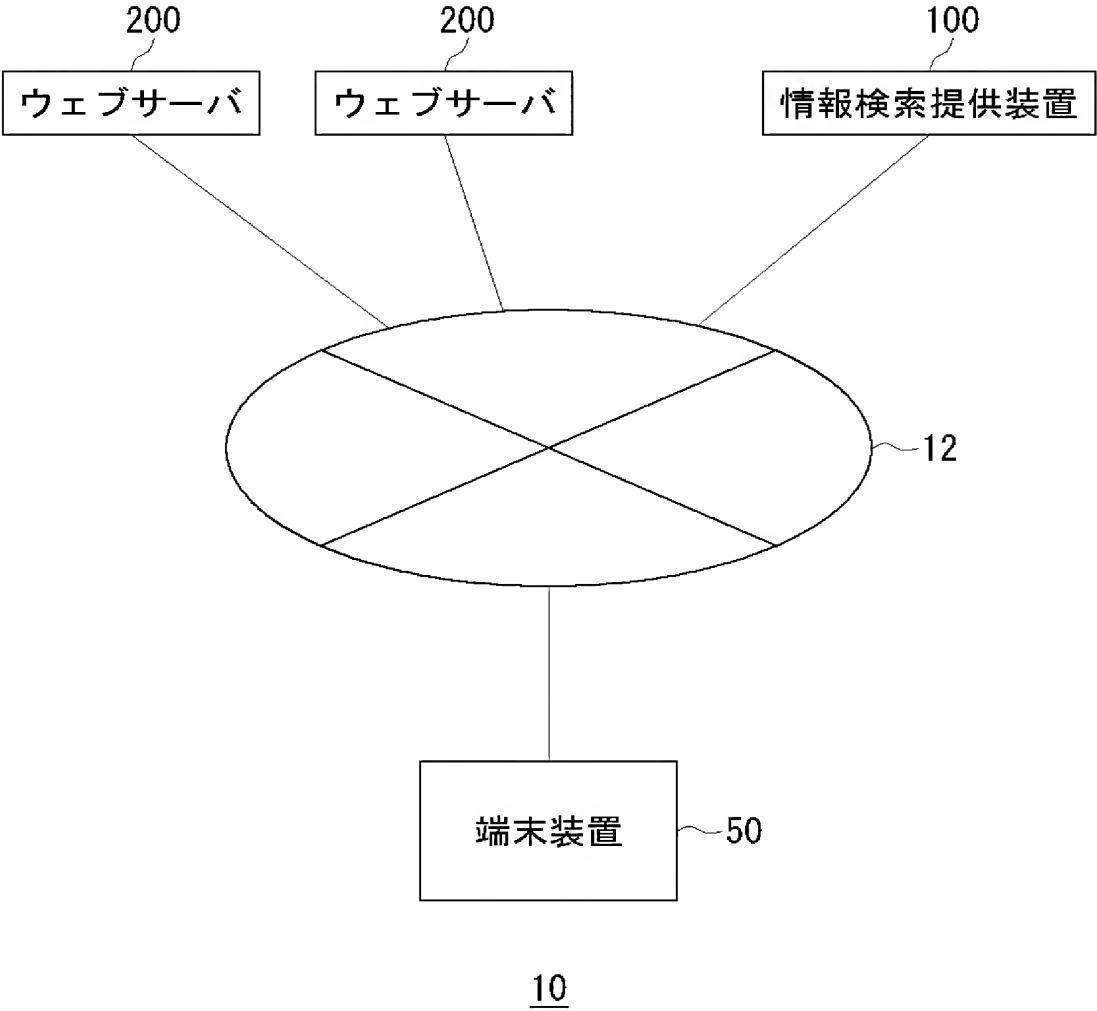
【図 19】検索初期画面の一例を示す図である。

【図 20】検索結果画面の一例を示す図である。

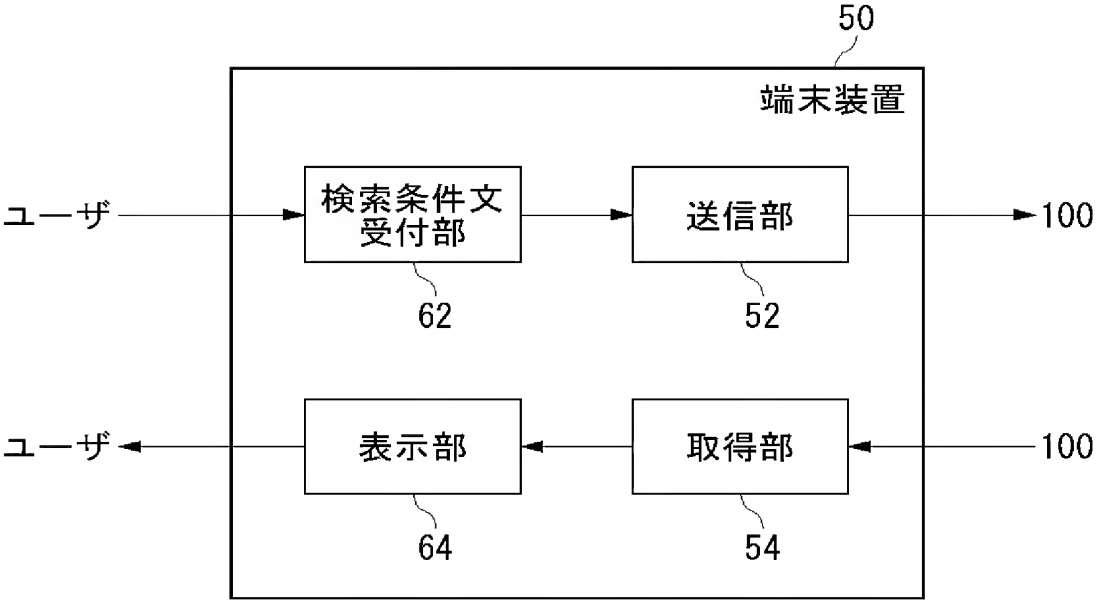
【符号の説明】

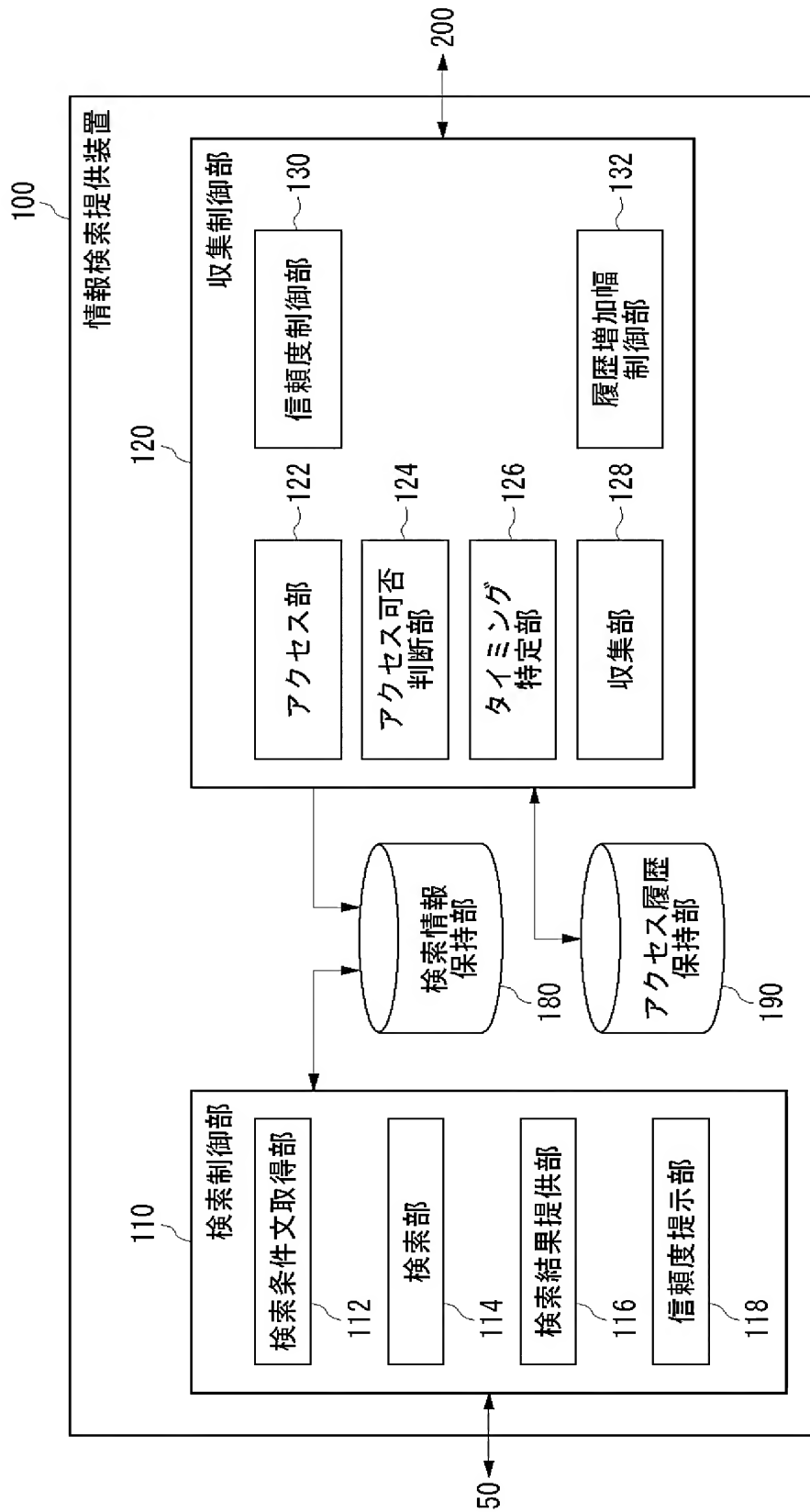
【0125】

10 情報検索提供システム、 12 ネットワーク、 50 端末装置、 54 取得部、 100 情報検索提供装置、 112 検索条件文取得部、 114 検索部、 122 アクセス部、 124 アクセス可否判断部、 126 タイミング特定部、 128 収集部、 130 信頼度制御部、 132 履歴増加幅制御部、 134 関連性判断部、 136 比較部、 138 更新判断部、 142 関連コンテンツ取得部、 144 要素抽出部、 146 分類部、 147 コンテンツ取得部、 148 自コンテンツ特定部、 150 整合性判断部、 190 アクセス履歴保持部、 200 ウェブサーバ、 R1 自コンテンツ。



【図 2】



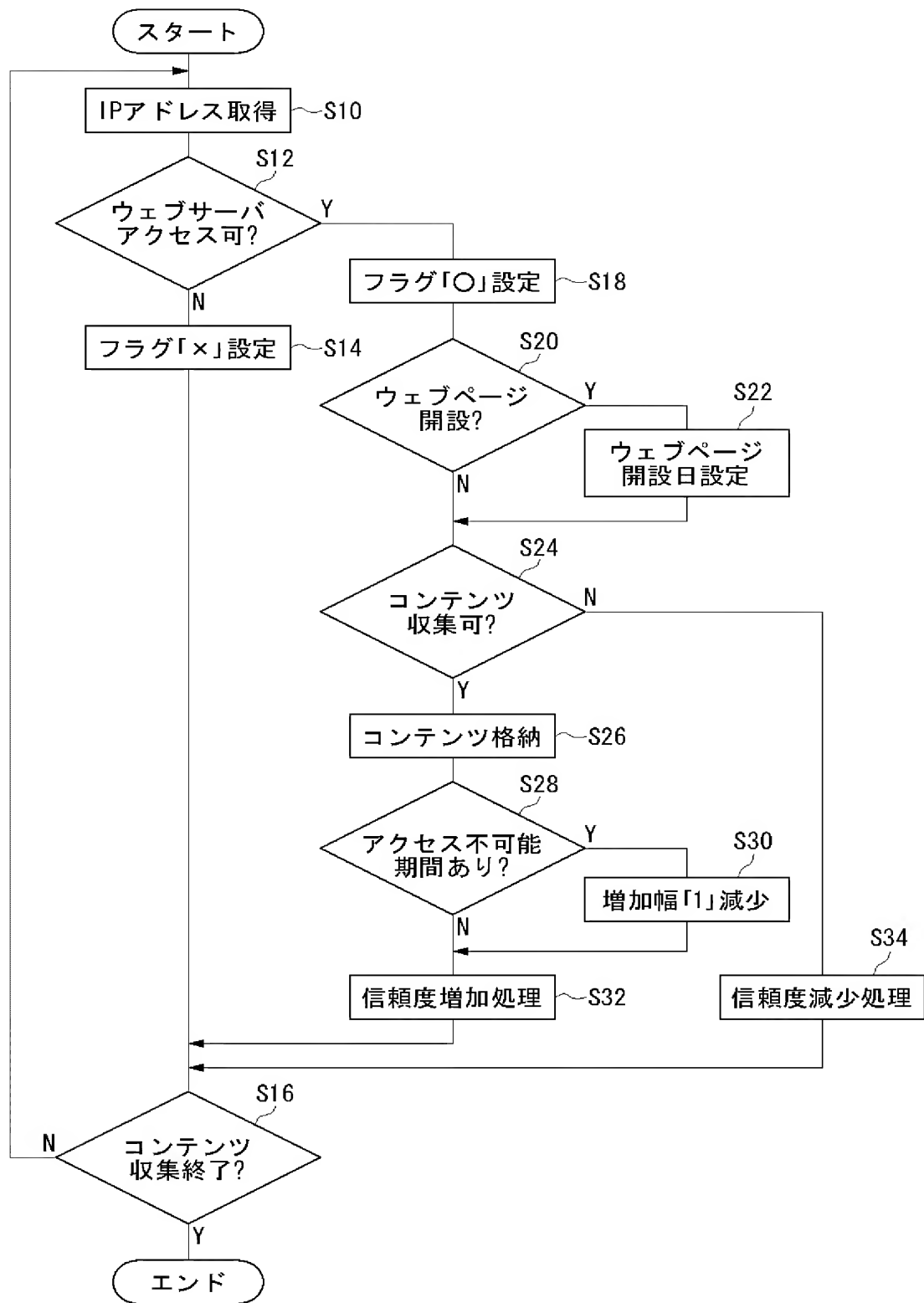


302 IPアドレス	304 URL	306 コンテンツ	308 ウェブページ 提供開始日	310 信頼度	312 増加幅
200.100.100.001	http://~1/	このサイトは鹿児島...	2004年1月1日	600	5
200.100.100.002	http://~2/	鹿児島テレビの...	2004年6月1日	80	5
200.100.100.003	http://~3/	鹿児島県の天気...	2004年5月3日	60	4

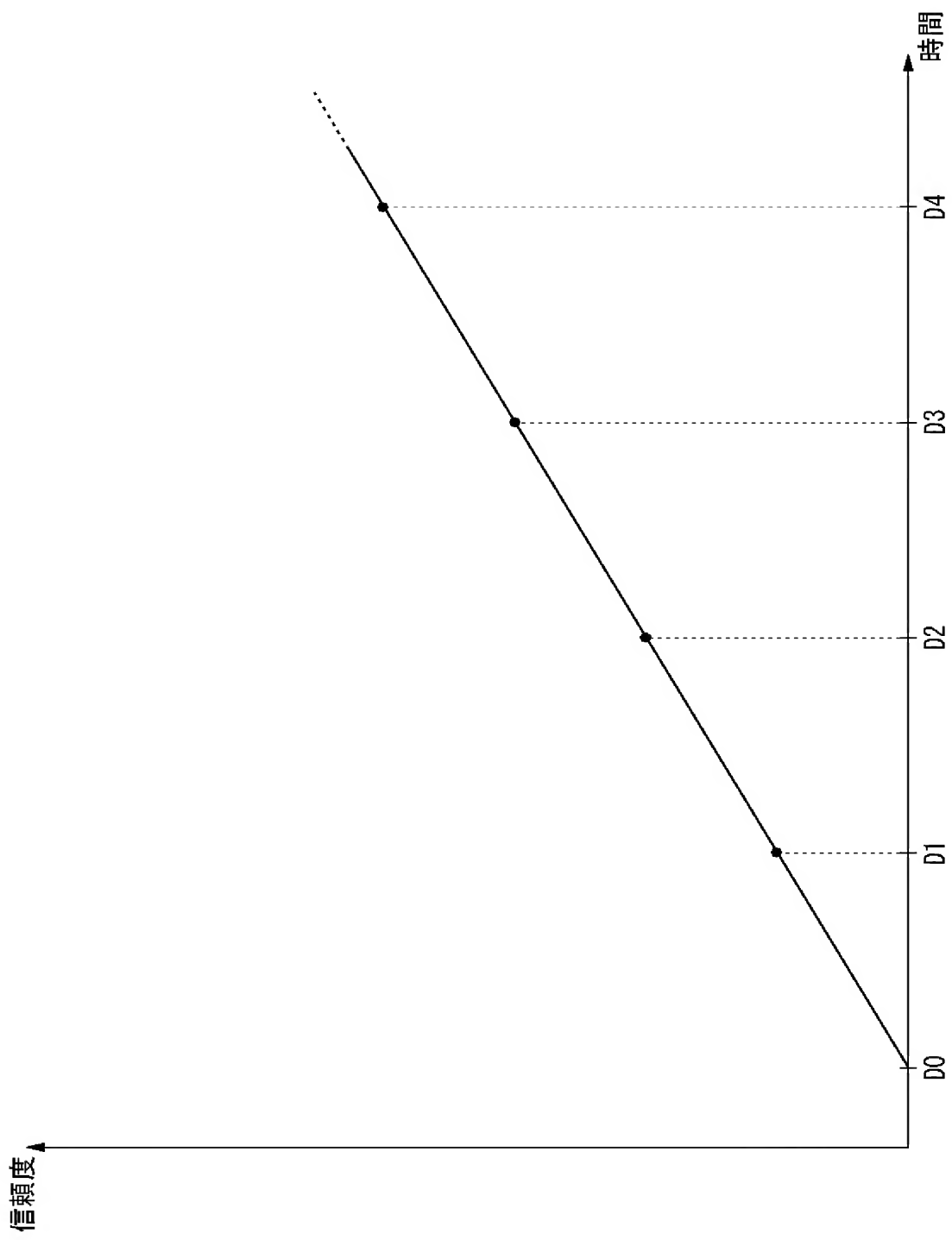
302

IPアドレス	アクセス履歴							322			
	2004年 1月1日	5月2日	5月3日	5月4日	5月31日	6月1日	6月2日				9月1日
〜	〜	〜	〜	〜	〜	〜	〜				〜
200.100.100.001	○	○	○	○	○	○	○				○ 324
200.100.100.002	×	×	×	×	×	○	○				○ 326
200.100.100.003	×	×	○	×	×	×	×				×
〜	〜	〜	〜	〜	〜	〜	〜				〜 328

【図 6】



【図 7】



信頼度

時間

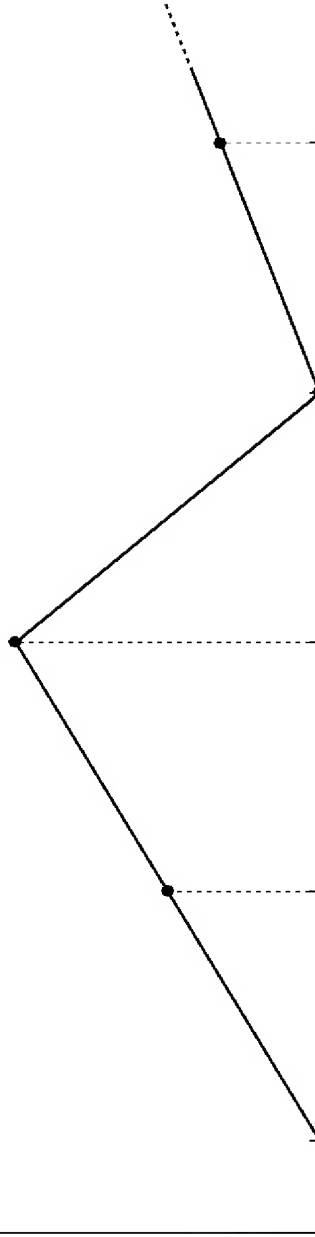
D0

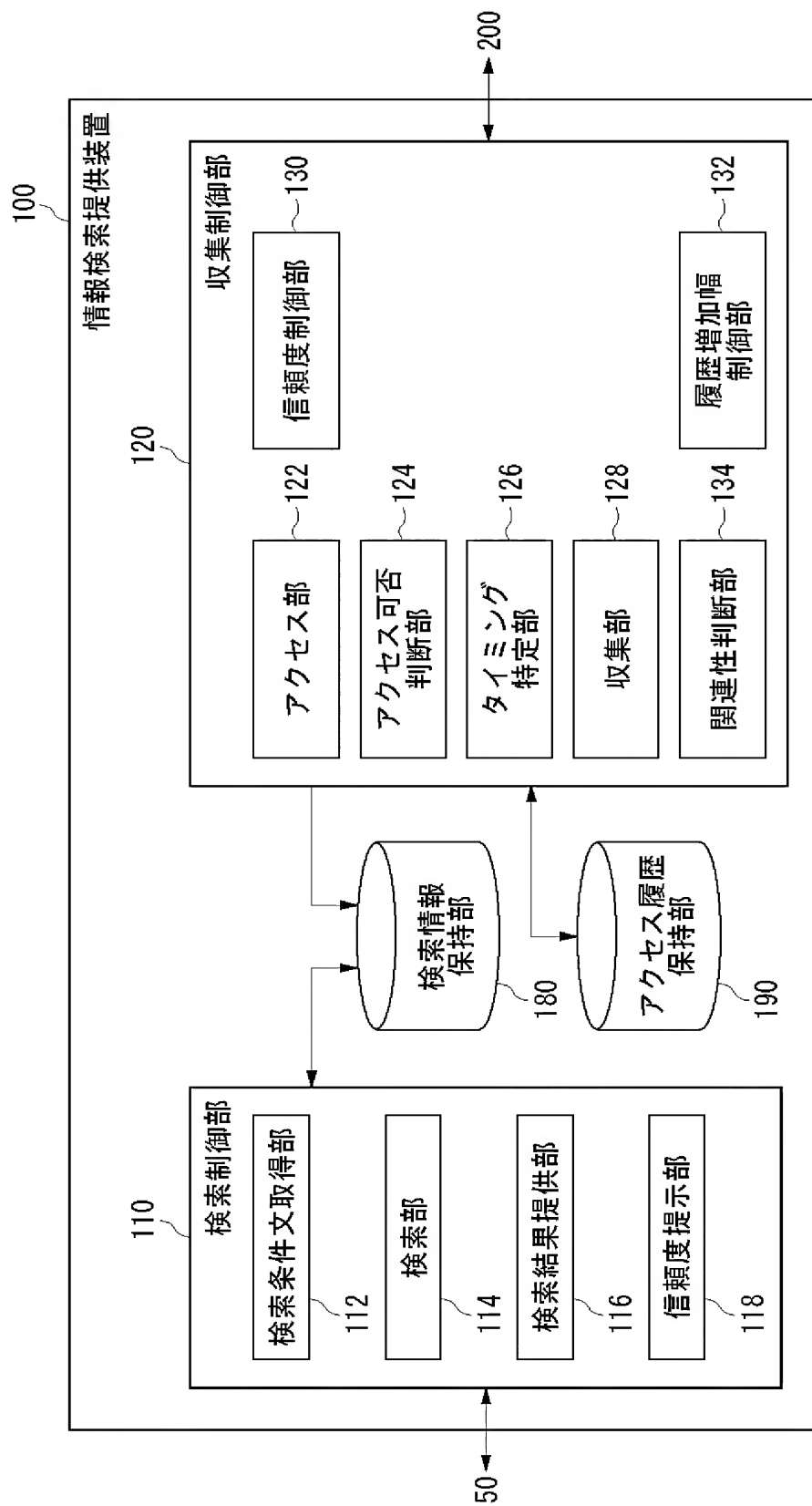
D1

D2

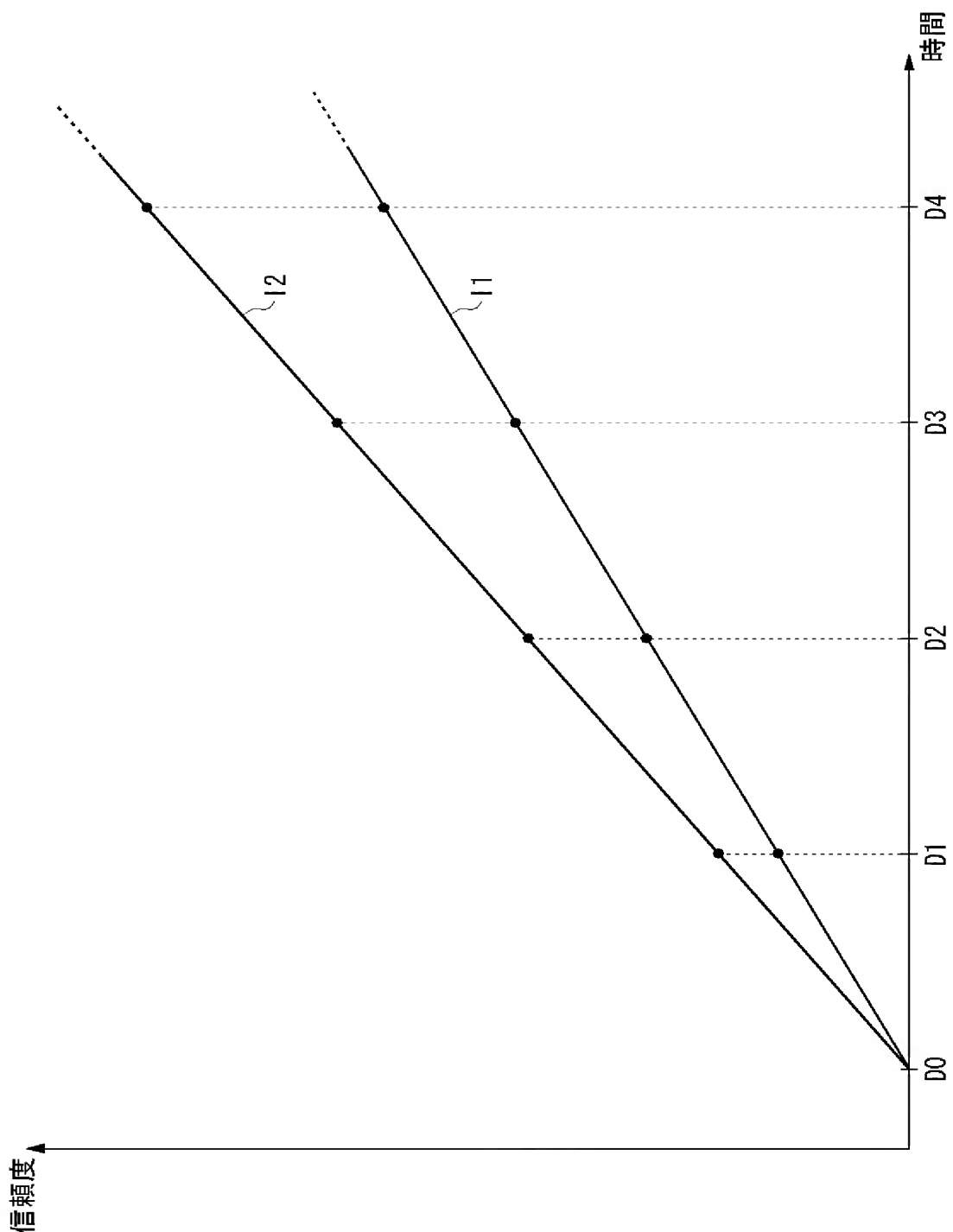
D3

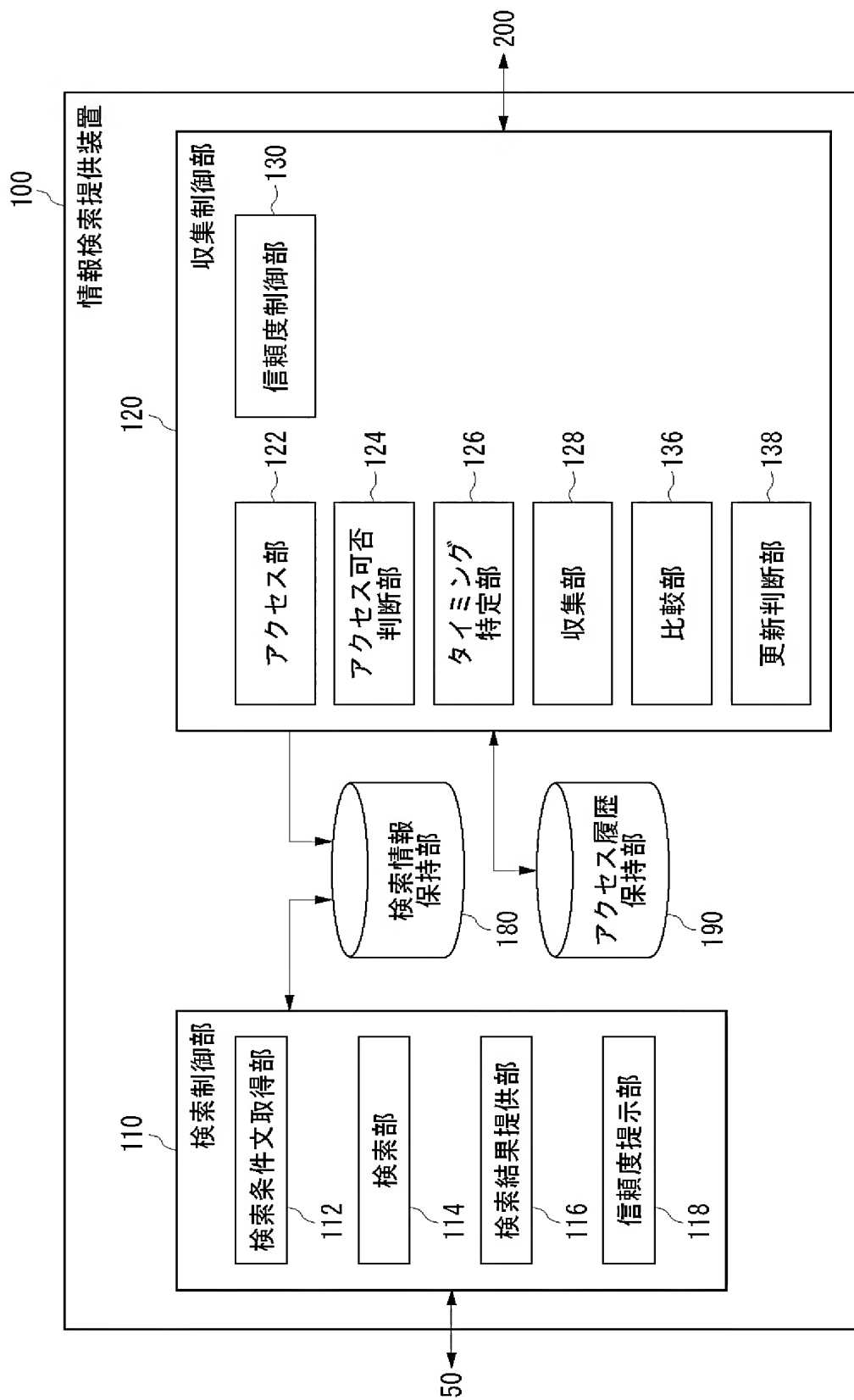
D4





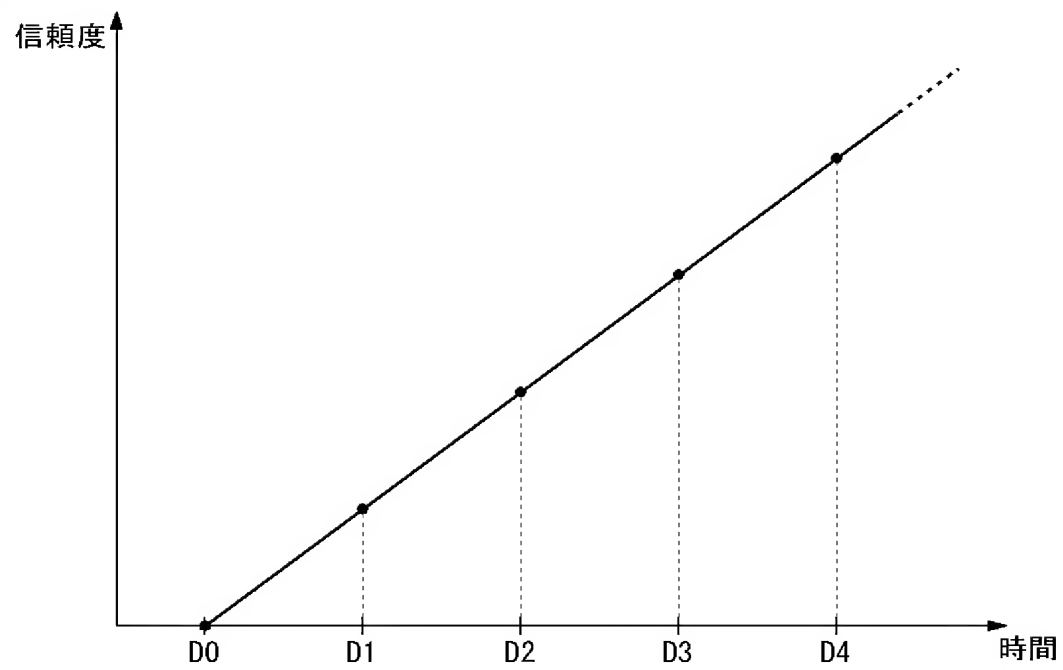
【図 1 0】



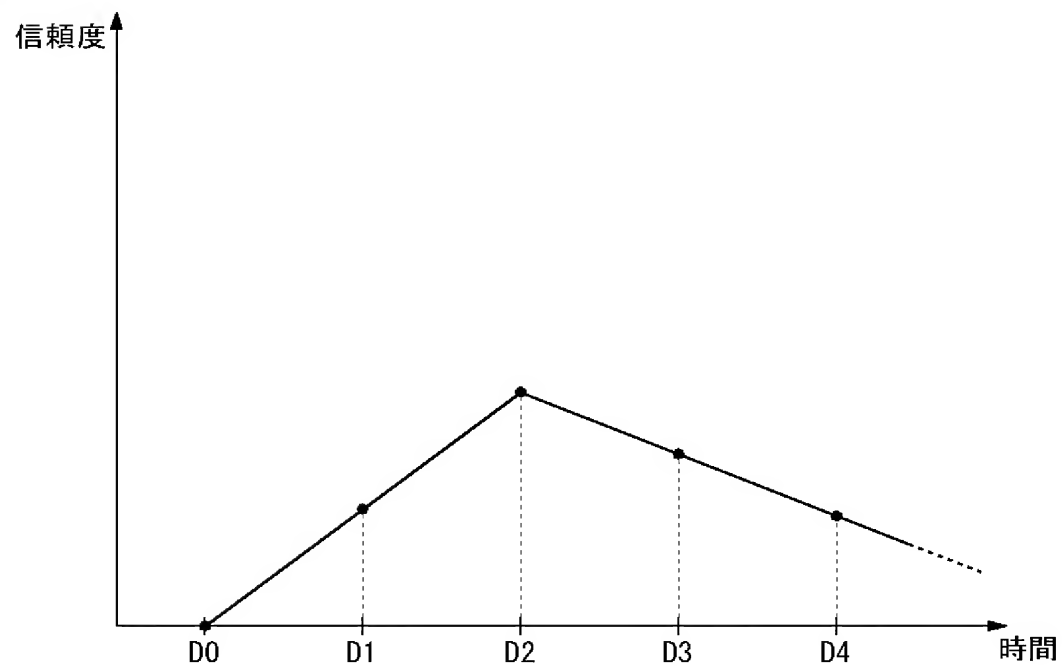


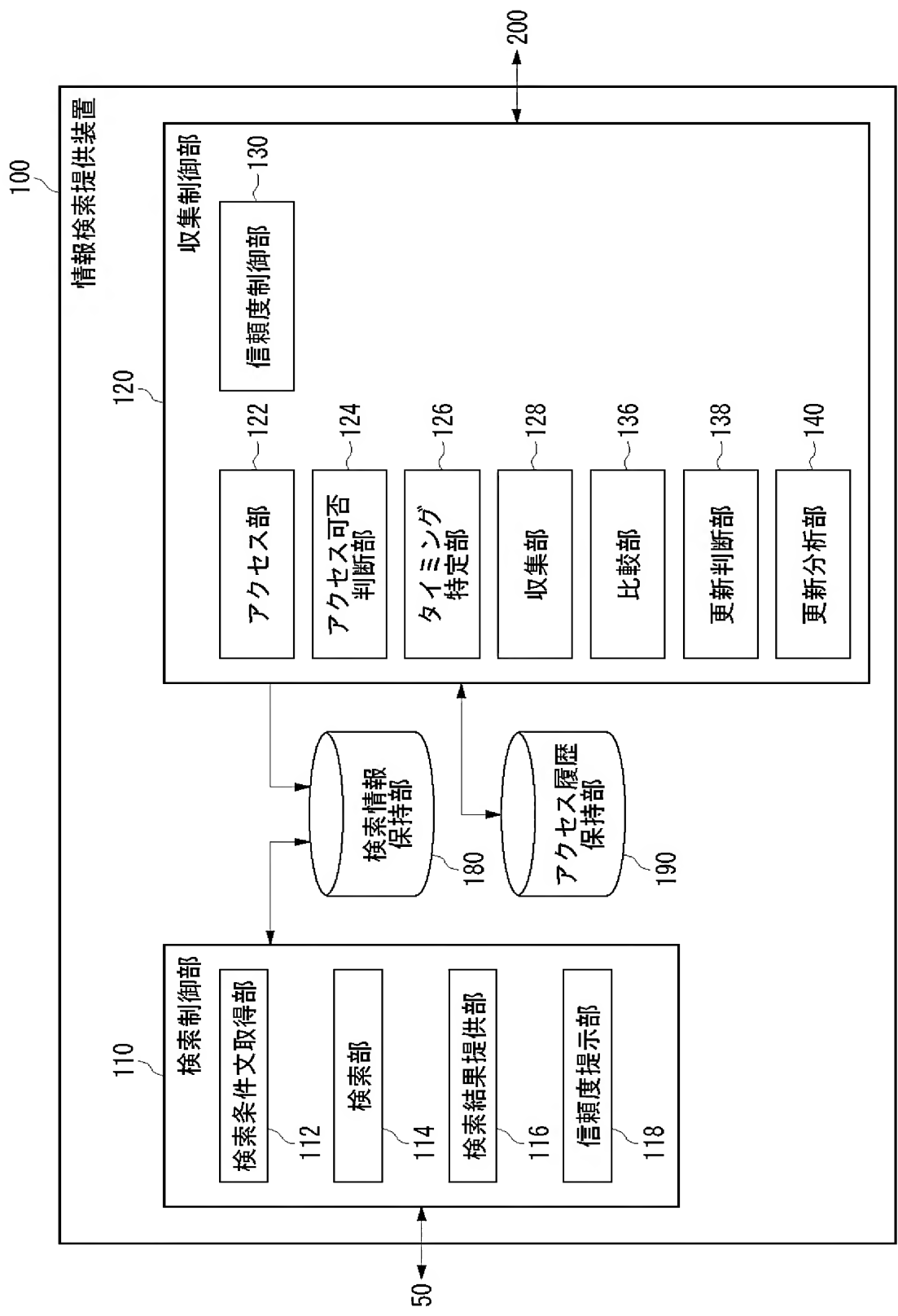
【図 1 2】

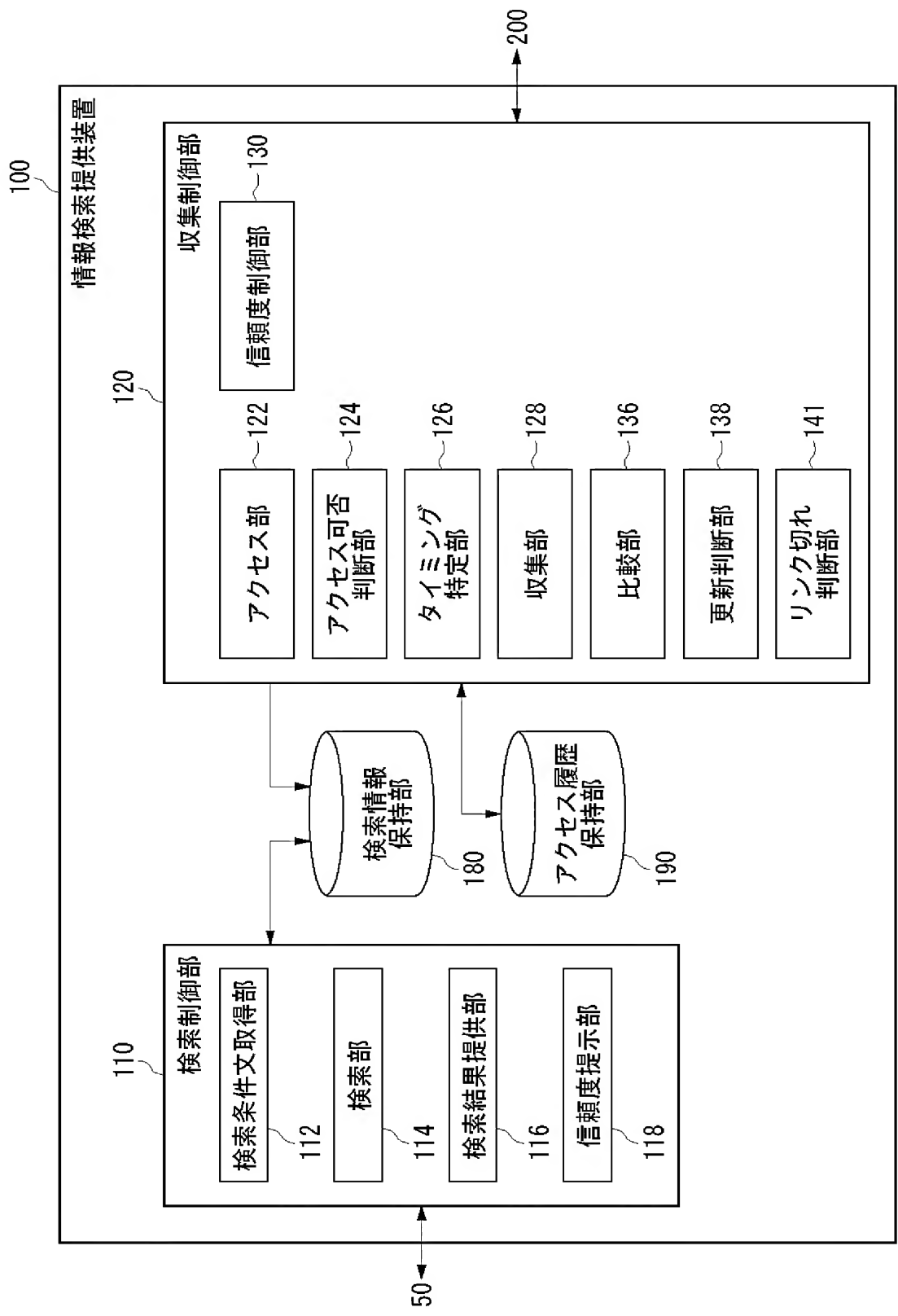
(a)



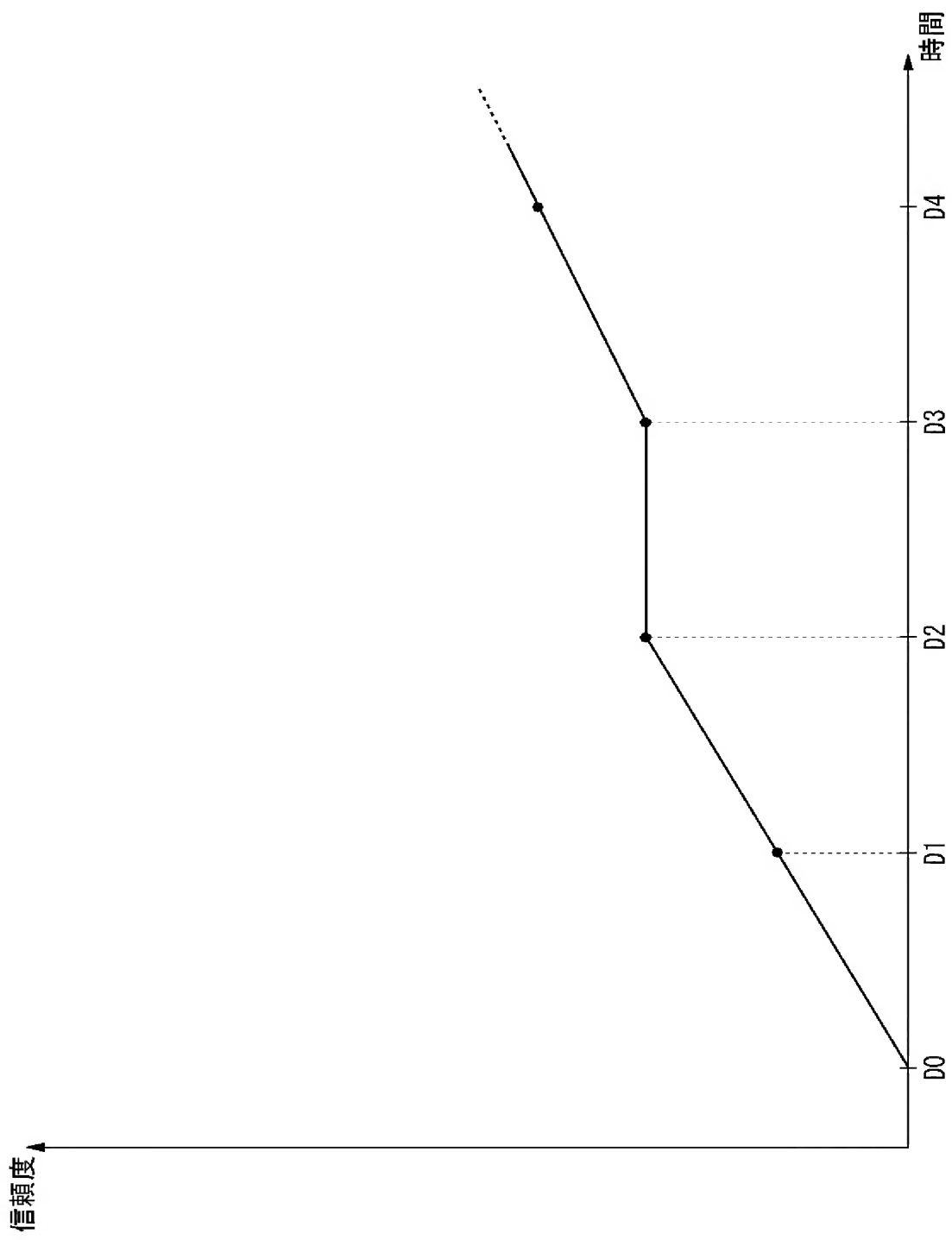
(b)

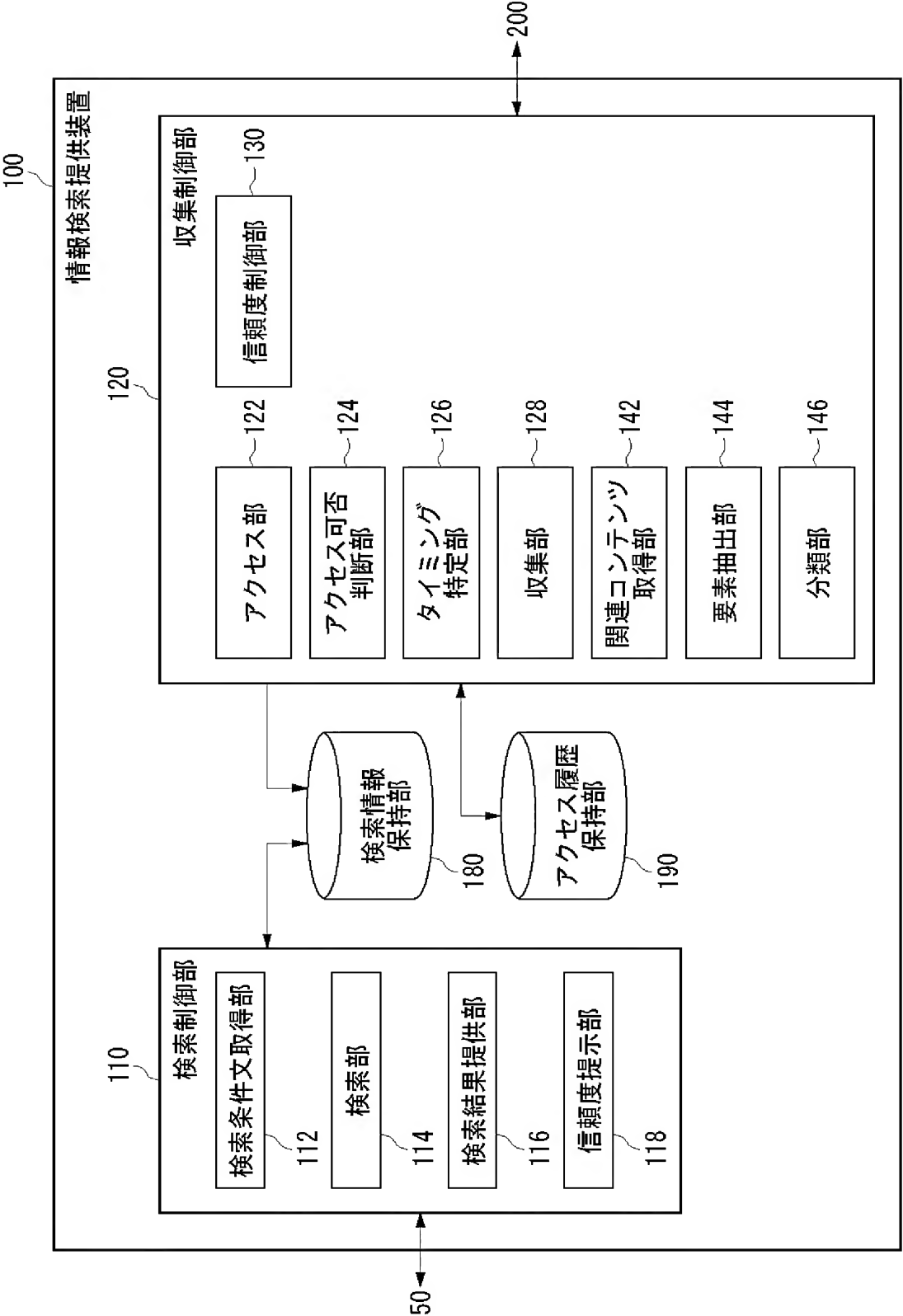


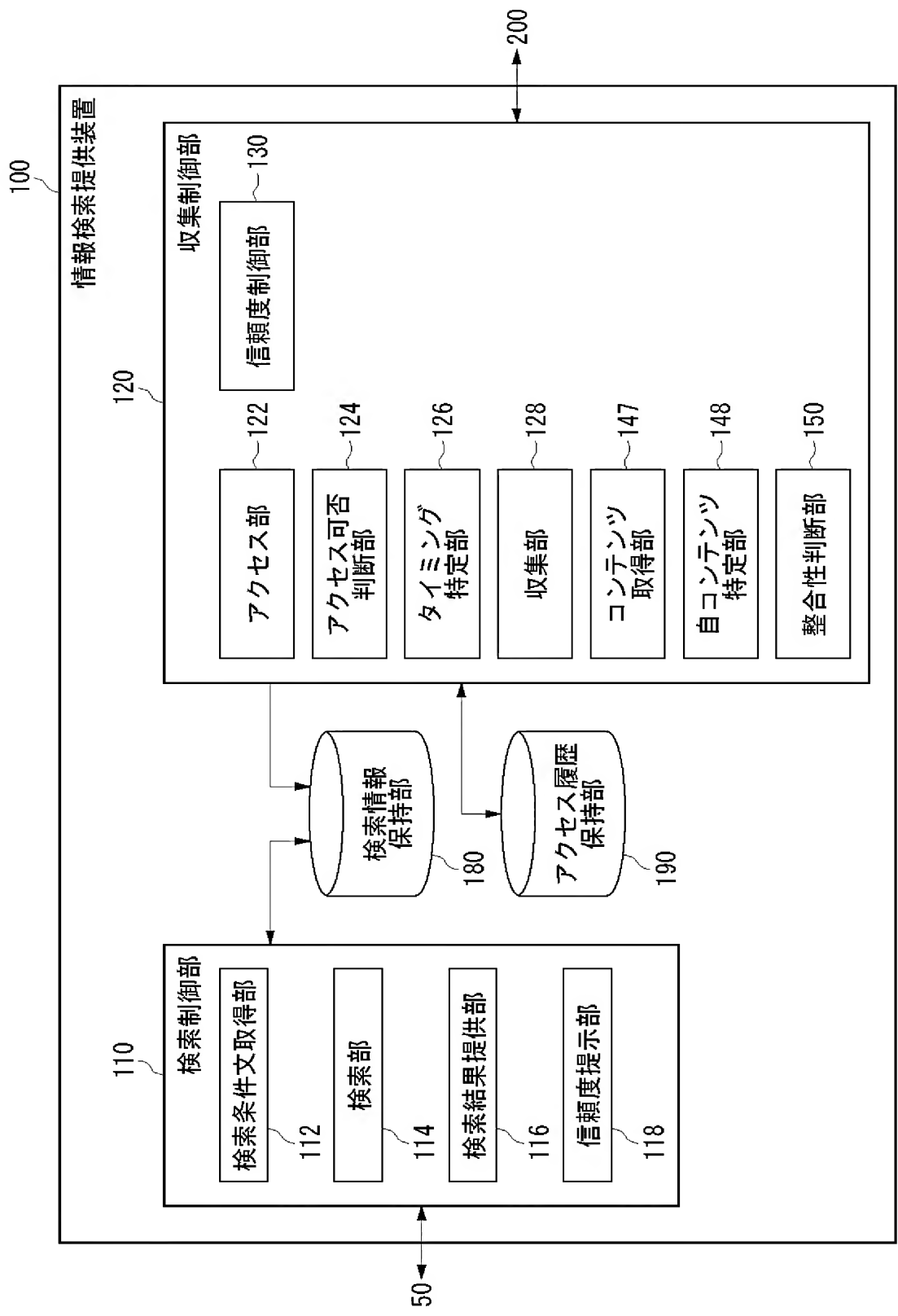




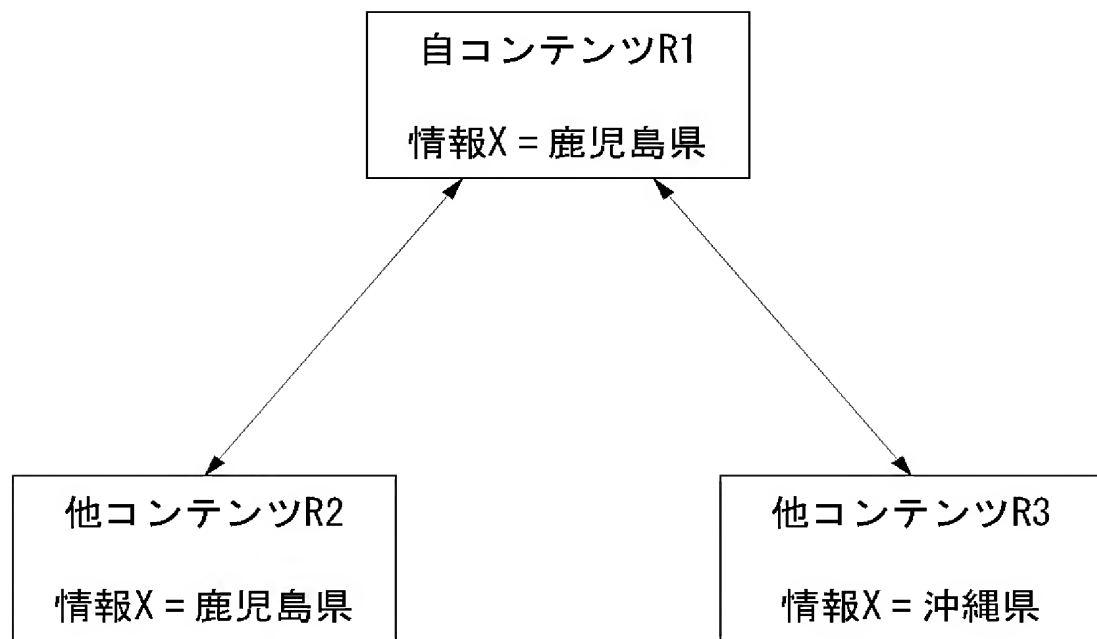
【図 1 5】







【図 1 8】



400

検索初期画面

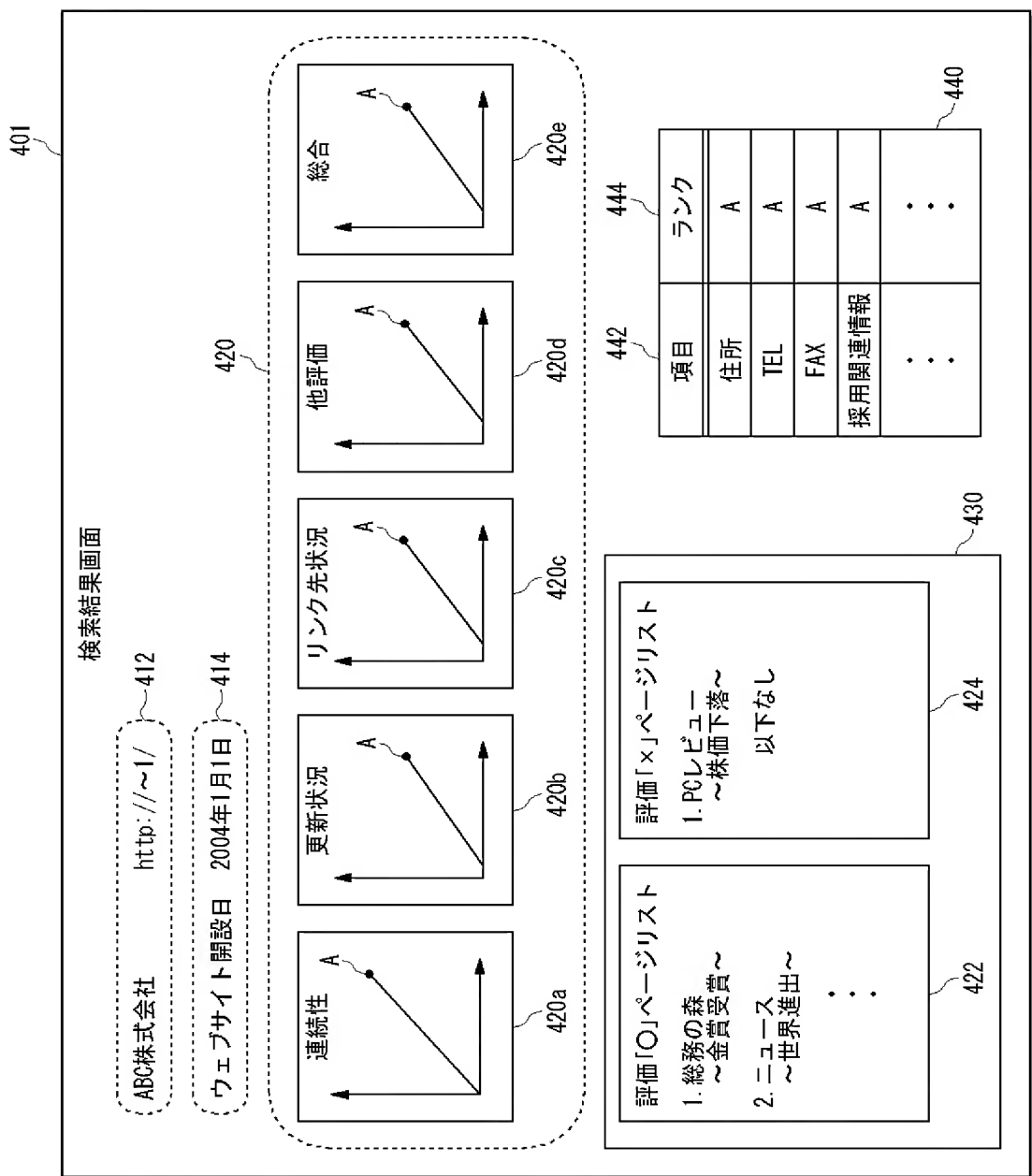
会社名/代表者名/商品名: 402
ABC株式会社

URL/TEL/FAX/住所: 404
http://~1/

その他: 406

408 開始

410 キャンセル

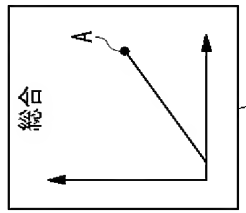
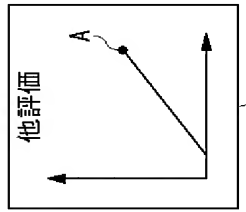
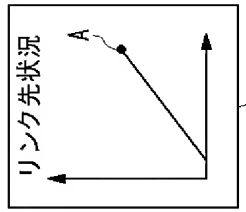
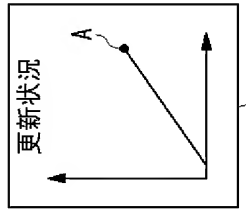
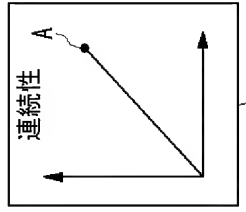


401

検索結果画面

ABC株式会社 http://~1/ 412

ウェブサイト開設日 2004年1月1日 414



評価「○」ページリスト

1. 総務の森
 ~金賞受賞~

2. ニュース
 ~世界進出~

 . . .

422

評価「×」ページリスト

1. PCレビュー
 ~株価下落~
 以下なし

424

442

項目	ランク
住所	A
TEL	A
FAX	A
採用関連情報	A
.

444

440

430

【書類名】 要約書

【要約】

【課題】 信頼のおけるコンテンツと信頼のおけないコンテンツとを選別する新たな情報検索サービスの仕組み作りが必要である。

【解決手段】 検索条件文を取得し当該検索条件文に合致したコンテンツを検索結果として提供する情報検索提供装置 1 0 0 であって、ウェブページに掲載される評価対象のコンテンツをネットワークを介して所定のタイミングで収集する収集部 1 2 8 と、収集部 1 2 8 による評価対象のコンテンツの収集の可否に応じて、評価対象のコンテンツの信頼度を増減せしめる信頼度制御部 1 3 0 と、を備える。これにより、評価対象のコンテンツの収集の可否に応じて当該コンテンツを信頼度の側面から評価することができ、その結果、信頼のおけるコンテンツと信頼のおけないコンテンツとを選別でき、ユーザ利便性の高い情報検索サービスを提供できる。

【選択図】 図 3

出願人履歴

5 9 6 1 7 7 5 5 9

20040722

名称変更

鹿児島市上荒田町 1 3 番 2 2 号

インターマン株式会社